

# Fuzzy Features Selection Technique for Brain MR Images

Meenakshi  
ECE, Hindu College of Engg.,  
Sonepat, India

R. B. Dubey  
ECE, Hindu College of Engg.,  
Sonepat, India

## ABSTRACT

Brain tumors are the second leading cause of cancer deaths in human throughout the world. Therefore accurate diagnosis is important for successful treatment of brain tumor. When modeling a complex, poorly defined problem with hundreds of possible inputs one must identify the significant inputs before any known modeling techniques can be applied. As generally the data contains many redundant features in it. Redundant features are those which provide no more information than the selected features and irrelevant features provide no more useful information in any context. Feature selection is the process in which one can select a subset of relevant features for use in model construction. Features selection techniques are very useful because they provide: improved model interpretability, shorter training times and enhanced generalization by reducing over fitting. For offering these benefits features selection techniques have become an apparent need in many bio-informatics applications such as brain MR images. In order to get best information a new fuzzy feature selection approach for MR brain tumor image have been used for the first time. The proposed method presents the two stages fuzzy curves and fuzzy surfaces. Fuzzy curves are used to isolate significant features from input and fuzzy surfaces to eliminate input features dependent on the significant input features. Results show that this method is able to achieve a very good reduction in the number of features.

## General Terms

Image Segmentation, Features Selection, Algorithm.

## Keywords

Fuzzy Curves, Fuzzy Surfaces, Features Selection, Image Segmentation.

## 1. INTRODUCTION

Brain tumors are the second leading cause of cancer-related deaths in human today. A brain tumor is a mass formed by an overgrowth of abnormal cells. Something triggers a single cell's genes to change, causing it to divide out of control. Brain tumors are often classified as benign or malignant. Truly benign tumors are composed of noncancerous cells that don't invade brain or other tissues. A malignant brain tumor may contain benign-appearing cells that invade normal tissue or contain cancerous cells either from the brain or other body cancers [1].

Feature selection is the process of choosing a small subset of features that ideally is necessary and sufficient to describe the target concept. Nowadays, the need to process large databases is becoming increasingly common. Furthermore, when faced with many noisy features, some algorithms take very long

time to converge, or never converge at all which will tend to construct poor classifiers.

Various introduced features during the training of a classifier are either partially or completely redundant to the target concept; an irrelevant feature does not affect the target concept at all, and a redundant feature does not add anything new to the target concept. In many applications, the size of a dataset is so large that learning might not work as well before removing these unwanted features. Reducing the number of redundant features drastically reduces the running time of a learning algorithm and gives a better concept. This helps in getting a better insight into the underlying concept of a real-world classification problem.

Various optimization approaches have been used for feature selection, such as exhaustive, heuristic and non-deterministic method. Statistical analysis method: sequential forward search (SFS) and sequential backward search (SBS) are widely used to select features [2-4]. The artificial neural network (ANN) and genetic algorithms (GA) has also been used which gives theoretically better results [2, 5, 6].

Previously different neural models were designed to select and extract the features of the MRI brain images; some of the related work is reported in [7-11]. The use of fuzzy systems has produced successful results in many areas, including pattern classification, optimization, and control of processes [12-13]. Sugeno and Yasukawa [14] propose a method to create a fuzzy model from input-output data. To identify the significant input variables, they build a one input fuzzy system for every input candidate. The input of the fuzzy system corresponding to the best performance index is chosen as the first significant input. Taking this input and every remaining input candidate, they build a series of two input fuzzy systems. The two inputs corresponding to the best performance measure are considered as two important inputs. They keep adding input variables one at a time until the performance index cannot be improved. Takagi and Hayashi [15] propose a neural network driven fuzzy reasoning system to analyze input-output data. They identify the input variables by eliminating each input variable and checking the performance index. J.B. Antoine et. al proposed a fuzzy surface approach for the registration of single photon emission computed tomography (SPECT) and MRI brain images. Fuzzy surface is a novel approach in which they used the morphological approach like morphological erosion and dilation [16]. Maria- Elena Algorri et. al classified the anatomical structures in MR brain images using fuzzy parameters [17]. Basically, fuzzy systems are systems with variables based on fuzzy logic. The rule based fuzzy systems (RBFS), focus of interest for this work, usually have two main components: a knowledge base (KB) and an inference mechanism (IM). There are several approaches for the

automatic definition of data bases and automatic generation of knowledge bases from data, including clustering algorithms, neural networks, and Genetic Algorithms (GA), which are among the well-succeeded ones [18-22]. There has been a considerable research effort focussed on the use of GA in the design of fuzzy systems. This initiative coined the term Genetic Fuzzy Systems (GFS), which are basically fuzzy systems with a learning process controlled by GA [21, 23]. Among the many GFS proposed approaches, a promising one uses GA to generate FRBs based on previously defined and fixed fuzzy sets [24]. This approach is adopted in [25], where a method consisting of three consecutive steps is proposed: (i) a feature selection process, (ii) the use of a GA to induce rules, and (iii) the use of another GA to eliminate unnecessary rules. The Fuzzy-wrapper method for pre-selection of features uses the Wang and Mendel method to generate FRBs as the base for a wrapper, is the best-first heuristic search technique [27].

To solve existing problems, we proposed a new fuzzy feature selection technique. The rest of the paper is organized as follows. Feature extraction is described in Section 2. Section 3 describes the proposed methodology. Section 4 gives results and discussions. Section 5 concludes the major findings of this research work and discusses the potential for further research.

## 2. FEATURES EXTRACTION

Texture is a very important characteristics used in identifying objects or regions of interest in an image. Texture contains important information about the structural arrangement of surfaces. The textural features based on gray-tone spatial dependencies have a general applicability in image classification [28].

### 2.1 Gray Level Co-occurrence Matrix

GLCM is defined as a two dimensional histogram of gray levels for a pair of pixels, which are separated by a fixed spatial relationship. GLCM of an image is computed using a displacement vector  $d$ , defined by its radius  $a$  and orientation  $e$ . Consider an image represented by a four grey-tone values 0 through 3. A generalized GLCM the image is shown here tones  $i$  and  $j$  have been neighbors satisfying the condition stated by displacement vector  $d$ .

The four GLCM for angles equal to  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$  and radius equal to 1. These are symmetric matrices hence evaluation of either upper or lower triangle serves the purpose. Frequency normalization can be employed by dividing value in each cell by the total number of pixel pairs possible. Hence the normalization factor for  $0^\circ$  would be  $(N_x - 1) \times N_y$  where  $N_x$  represents the width and  $N_y$  represents the height of the image. The quantization level is an equally important consideration for determining the co-occurrence texture features. Also, neighboring co-occurrence matrix elements are highly correlated as they are measures of similar image qualities. In this work, 21 GLCMs are constructed in order to compute 4 angles (0, 45, 90, 135), each with unit pixel distances of 1, 3 and 5 respectively. The 22 features described above are computed for the 21 GLCMs, thus, resulting in 21 texture features for each image.

#### 2.1.1 Choice of radius $a$

From the previous studies, it has been concluded that overall classification accuracies with  $a = 1, 2, 4, 8$  are acceptable with the best results for  $a = 1$  and 2. This conclusion is justified, as a pixel is more likely to be correlated to other closely located pixel than the one located far away. Also, displacement value

equal to the size of the texture element improves classification.

#### 2.1.2 Choice of angle $e$

Every pixel has eight neighbouring pixels allowing eight choices for  $e$ , which are  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ ,  $180^\circ$ ,  $225^\circ$ ,  $270^\circ$  or  $315^\circ$ . However, taking into consideration the definition of GLCM, the co-occurring pairs obtained by choosing  $e$  equal to  $0^\circ$  would be similar to those obtained by choosing  $e$  equal to  $180^\circ$ . This concept extends to  $45^\circ$ ,  $90^\circ$  and  $135^\circ$  as well. Hence, one has four choices to select the value of  $e$ . Sometimes, when the image is isotropic, or directional information is not required, one can obtain isotropic GLCM by integration over all angles.

#### 2.1.3 Choice of quantized gray level ( $G$ )

The dimension of a GLCM is determined by the maximum gray value of the pixel. Number of gray levels is an important factor in GLCM computation. More levels would mean more accurate extracted textural information, with increased computational costs. The computational complexity of GLCM method is highly sensitive to the number of gray levels and is proportional to  $O(G^2)$ , where  $O(G^2)$  is time complexity and  $G$  is the number of gray levels in an image [29, 30].

Thus for a predetermined value of  $G$ , a GLCM is required for each unique pair of  $a$  and  $e$ . GLCM is a second-order texture measure. The GLCM's lower left triangular matrix is always a reflection of the upper right triangular matrix and the diagonal always contains even numbers. Various GLCM parameters are related to specific first-order statistical concepts. For instance, contrast would mean pixel pair repetition rate, variance would mean spatial frequency detection etc. Association of a textural meaning to each of these parameters is very critical. Traditionally, GLCM is dimensioned to the number of gray levels  $G$  and stores the co-occurrence probabilities  $p_{ij}$ . To determine the texture features, selected statistics are applied to each GLCM by iterating through the entire matrix. The textural features are based on statistics which summarize the relative frequency distribution which describes how often one gray tone will appear in a specified spatial relationship to another gray tone on the image.

The question what exactly the textural features represent from a human perception point of view can be a subject for a thorough experimentation. The textural features described above, the angular second moment, the entropy, the sum entropy, the difference entropy, the information measure of correlation and the maximal correlation features have the invariance property. Nineteen features have been extracted from GLCM [31].

**Table 1: Various texture features**

Features	
1. Autocorrelation	$f_1 = \sum_i \sum_j (ij) p(i, j)$
2. Contrast	$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)  i - j  = n \right\}$
3. Correlation	$f_3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$

4. Cluster Prominence	$f_4 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^4 p(i, j)$
5. Cluster Shade	$f_5 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^3 p(i, j)$
6. Dissimilarity	$f_6 = \sum_i \sum_j  i - j  p(i, j)$
7. Energy	$f_7 = \sum_i \sum_j p(i, j)^2$
8. Entropy	$f_8 = -\sum_i \sum_j p(i, j) \log(p(i, j))$
9. Homogeneity	$f_9 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j)$
10. Maximum Probability	$f_{10} = \max(p(i, j))$
11. Variance	$f_{11} = \sum_i \sum_j (i - j)^2 p(i, j)$
12. Sum Average	$f_{12} = \sum_{i=2}^{2N_g} i p_{x+y}(i)$
13. Sum Variance	$f_{13} = \sum_{i=2}^{2N_g} (i - f_{14})^2 p_{x+y}(i)$
14. Sum Entropy	$f_{14} = -\sum_{i=2}^{2N_g} p_{x+y}(i) \log(p_{x+y}(i))$
15. Inverse Difference Normalized (INN)	$f_{15} = \sum_{i,j=1}^G \frac{p_{x+y}(i, j)}{1 +  i - j ^2 / G^2}$
16. Difference Entropy	$f_{16} = -\sum_{i=0}^{N_g-1} p_{x-y}(i) \log(p_{x-y}(i))$
17. Information Measure of Correlation-1	$f_{17} = \frac{\sum_i \sum_j p(i, j) (\log(p(i, j)) - \log(p_x(i) p_y(i)))}{\max(HX, HY)}$
18. Information Measure of Correlation-2	$f_{18} = \sqrt{1 - e^{-2(a-b)}}$ $a = -\sum_i \sum_j p_x(i) p_y(i) \log(p_x(i) p_y(i))$

$b = -\sum_i \sum_j p(i, j) \log(p_x(i) p_y(i))$
19. Inverse Difference Moment Normalized
$f_{19} = \sum_{i,j=1}^G \frac{p_{x+y}(i, j)}{1 + (i - j)^2 / G^2}$

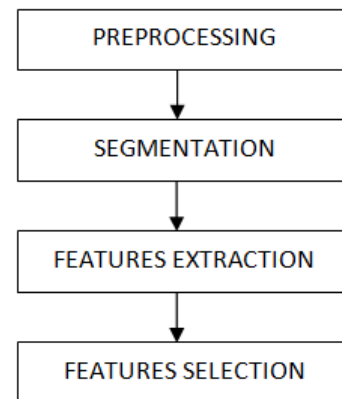
### 3. PROPOSED METHODOLOGY

Here image is filtered before segmentation and then features are extracted. Finally feature selection is carried out using fuzzy curves and fuzzy surfaces respectively. The fuzzy inference system used for this features selection purpose is Mamdani fuzzy inference.

When the input data to an algorithm is too large to be processed and it is suspected to be redundant then the input data will be transformed into a reduced representation set of features. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. To address the problem of classification MRI brain images by providing selected features set so that a robust and more accurate classifier can be designed [26].

When performing analysis of complex data, one of the major problems stems from the number of variables is involved [32]. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. The procedure of feature selection is as follows: in accordance with pre-designed selection criteria, the most important features of the given input data are selected by the optimal operations under the prefixed criterion and the remaining features are removed from the input to reduce the data amount [15].

The proposed methodology consists of four main steps: preprocessing, segmentation, feature extraction and feature selection.

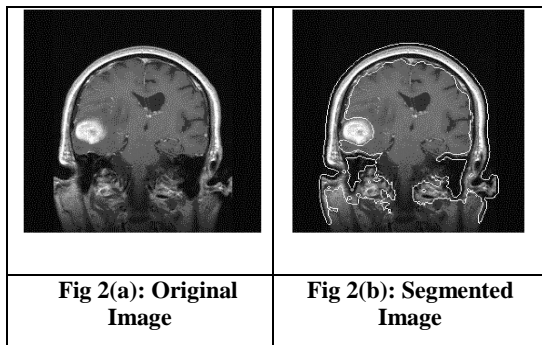


**Fig 1: Proposed Methodology**

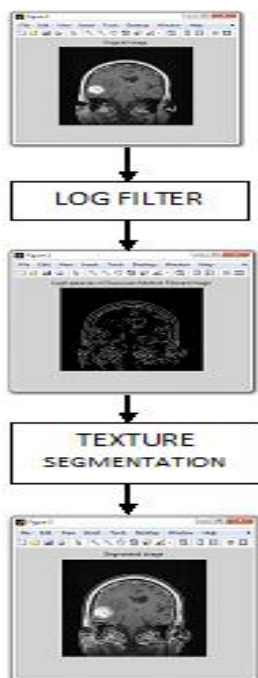
### 3.1 Preprocessing and Segmentation

Digital images are prone to a variety of types of noise. Noise is the result of errors in the image acquisition process that result in pixel values that do not reflect the true intensities of the real scene. There are several ways that noise can be introduced into an image, depending on how the image is created.

Texture segmentation has been used for segmenting the brain MR images. It is the identification of regions based on their texture. In texture segmentation first step is reading then create texture image. After obtaining texture image create rough mask for the bottom texture and use rough mask to segment the top texture and then image is segmented.



A pre-processing, usually noise-reducing step is applied to improve image quality. For preprocessing Laplacian of Gaussian filter (LoG) and texture segmentation is used before extracting features from images. Thus by providing image at input of LoG filter result obtained are as shown in Fig. 3.



**Fig 3: Preprocessing of images using LOG filter and texture segmentation**

### 3.2 Features Extraction

On segmented image using GLCM nineteen texture features are obtained. Texture features of a image are as shown in the Table 2. Similarly, twenty images are used from our data base to extract features.

**Table 2: Extracted features from brain MR image**

Extracted Features		Abb.
1.	Autocorrelation	autoc
2.	Contrast	contr
3.	Correlation-m	corrmm
4.	Cluster Prominence	cprom
5.	Cluster Shade	cshad
6.	Dissimilarity	dissi
7.	Energy	energ
8.	Entropy	entro
9.	Homogeneity-m	homomm
10.	Maximum probability	maxpr
11.	Variance	sosvh
12.	Sum average	savgh
13.	Sum variance	svarh
14.	Sum entropy	senh
15.	Difference entropy	denth
16.	Information measure of correlation1	inf1h
17.	Information measure of correlation2	inf2h
18.	Inverse difference normalized(INN)	indnc
19.	Inverse difference moment normalized	idmnc

### 3.3 Features Selection

After extracting all the features of an image features selection can be carried out. In this work features are selected using fuzzy curves and fuzzy surfaces. Fuzzy curves have taken one step to select fifteen features out of nineteen and then fuzzy surface is used. In using fuzzy surfaces features set selected by fuzzy curves are used and this step is repeated thrice to get the best features set from the complete set of extracted features [33].

#### 3.3.1 Features Selection using fuzzy curves

Fuzzy curves are based on an idea that the most important inputs do the best job of approximating the output [32]. It is a nonlinear continues curve, which establishes a connection between a specific input and the output, performing a projection of the multidimensional input and output space on the probed input-output space. The height of the projected output is the measure of importance of the specific input [34].

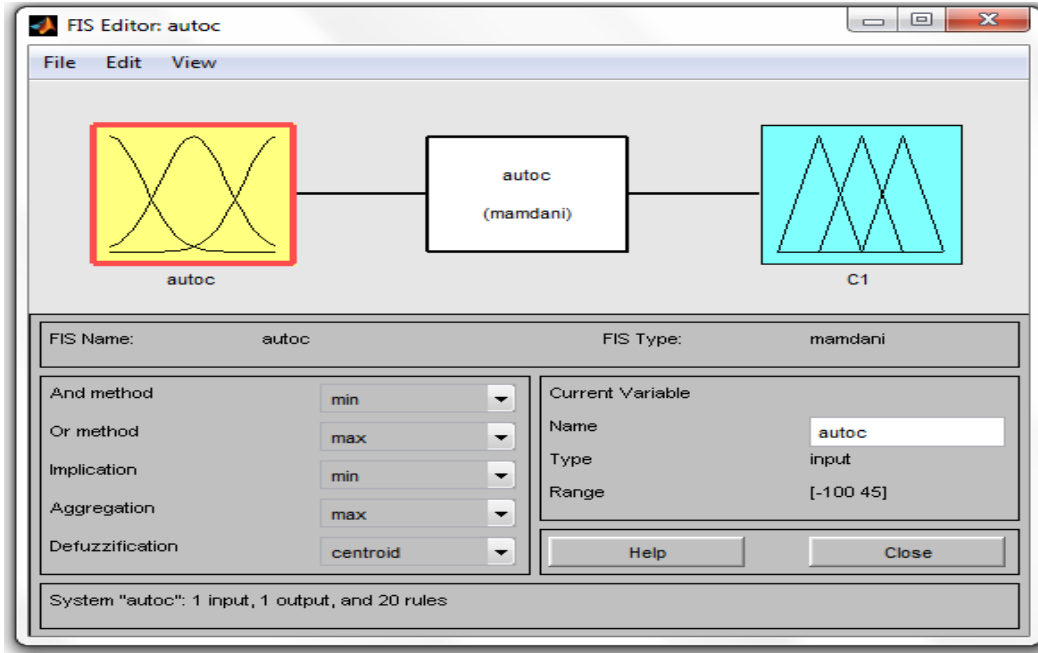


Fig 4: FIS Editor

Assuming there exists a feature vector consisting of  $N$ -features (inputs), belonging to a certain class  $y_k$  (single output), the behavior of a classifier for the classification of such vectors can be statistically represented by a set of  $m$  input/output observation data points  $(x_{i,k}, y_k)$ ,  $k = 1, 2, \dots, m$ . For each data point, we can create a fuzzy membership function  $\mu_{i,k}(x_i)$  using Eqn. (1) from the fuzzy rule: IF  $x_i$  is  $\mu_{i,k}(x_i)$ , THEN is  $y_k$ ,

that can be thought of as fuzzy rule for the output  $y$  with respect to each feature variable  $x_i$ ,

$$\mu_{i,k}(x_i) = \exp \left( - \left( \frac{x_{i,k} - x_i}{\sigma_i} \right)^2 \right) \dots \dots (1)$$

Each Gauss function is located at point  $(x_{i,k}, y_k)$  the parameter  $\sigma_i$  has a fixed value per feature variable  $x_i$ , which equals 5% ~ 20% of the  $x_i$  variable range [33].

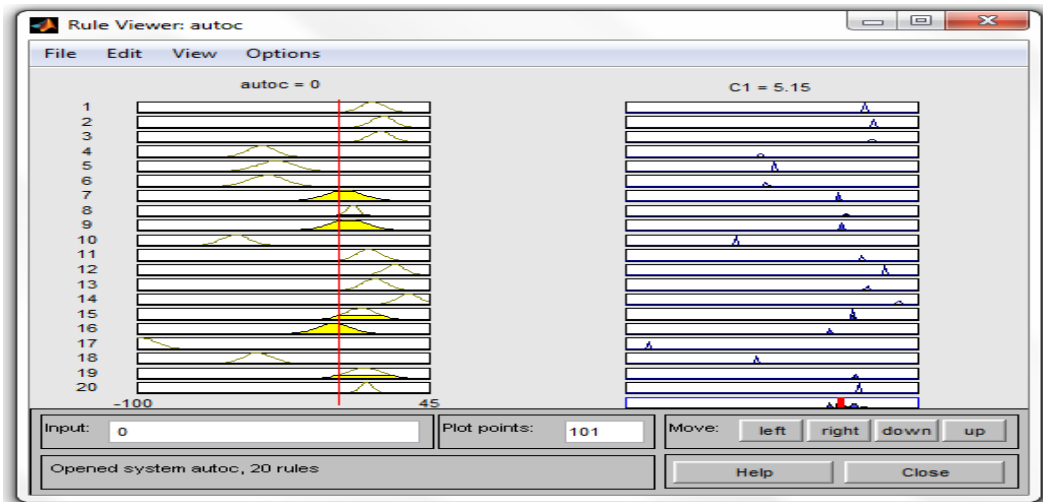


Fig 5: FIS Rule Viewer

For a feature variable, a fuzzy membership function base comprising  $m$  fuzzy membership functions is used. Using the centroid defuzzification technique, defuzzification of these fuzzy membership functions is done to create a fuzzy curve  $C_i$  for each feature variable  $x_i$  using the following formula:

$$C_{i,k}(x_i) = \frac{\sum_{k=1}^m y_k * \mu_{i,k}(x_i)}{\sum_{k=1}^m \mu_{i,k}(x_i)} \dots \dots (2)$$

Eqn. (2) provides a continuous curve, which approximates the input output data, and behaves as a fuzzy model. The mean

square error  $MSE_{C_i}$  is used to estimate the quality of the approximation,

$$MSE_{C_i} = \frac{1}{m} \sum_{k=1}^m (C_{i,k}(x_i) - y_k)^2 \dots \dots (3)$$

If  $MSE_{C_i}$  is large, the fuzzy curve  $c$  is doing a poor job of approximating the output; if the  $MSE_{C_i}$  is small, the fuzzy curve  $c$  is doing the best job of approximating the output [33].

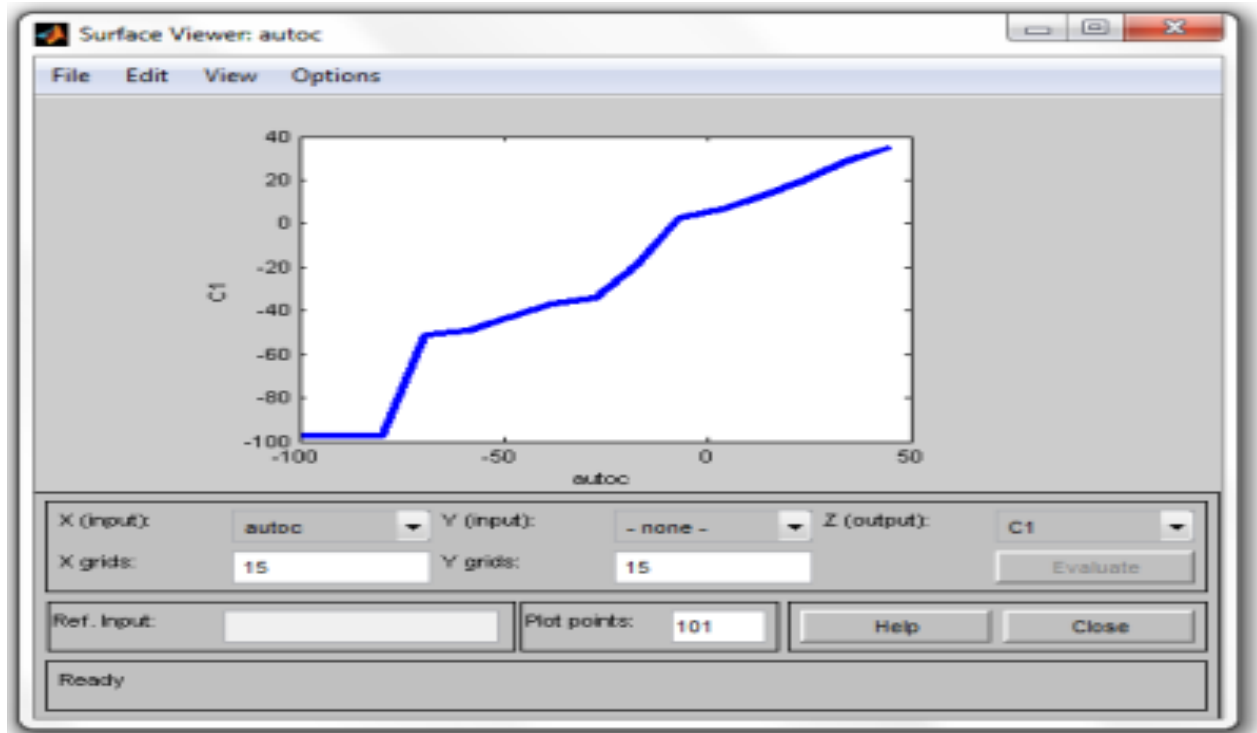


Fig 6: FIS Surface Viewer

Hence, compute the fuzzy curves  $C_i$  for all candidate feature variable  $x_i$  and compute the mean square error  $MSE_{C_i}$  for each fuzzy curve. Rank the feature variables in ascending  $MSE_{C_i}$  order. The feature with the smallest  $MSE_{C_i}$  is most important, the feature with the largest  $MSE_{C_i}$  is least important and the features (20% of input features) whose fuzzy curves have the largest  $MSE_{C_i}$  are eliminated as not significant. Using fuzzy curves, one can rapidly identify the important features from the set of candidate features. However, when features are dependent one cannot discard the dependent feature variables.

### 3.3.2 Features Selection using fuzzy surfaces

The fuzzy surface is an extension of the fuzzy curve, it is based on the simple idea: independent feature do better job of approximating the output than dependent inputs [33].

Use the most important feature and each of other remaining features, chosen by the fuzzy curve, to create fuzzy surfaces, rank the features in order of dependence with the most important feature (in descending surface values). The feature with the lowest mean square error is judged to be the second most important feature. The other features (20% of input features) with the largest mean square errors are probably related to the most important feature, and they are eliminated because they are the most closely dependent or related to the most important feature.

Fuzzy surface can be thought of as a two-dimensional fuzzy curve. Fuzzy surface is defined as [33]:

$$S_{i,j}(x_i, x_j) = \frac{\sum_{k=1}^m y_k * \mu_{i,k}(x_i) + \sum_{k=1}^m y_k * \mu_{j,k}(x_j)}{2 \sum_{k=1}^m \mu_{i,k}(x_i) * \sum_{k=1}^m \mu_{j,k}(x_j)} \dots \dots (4)$$

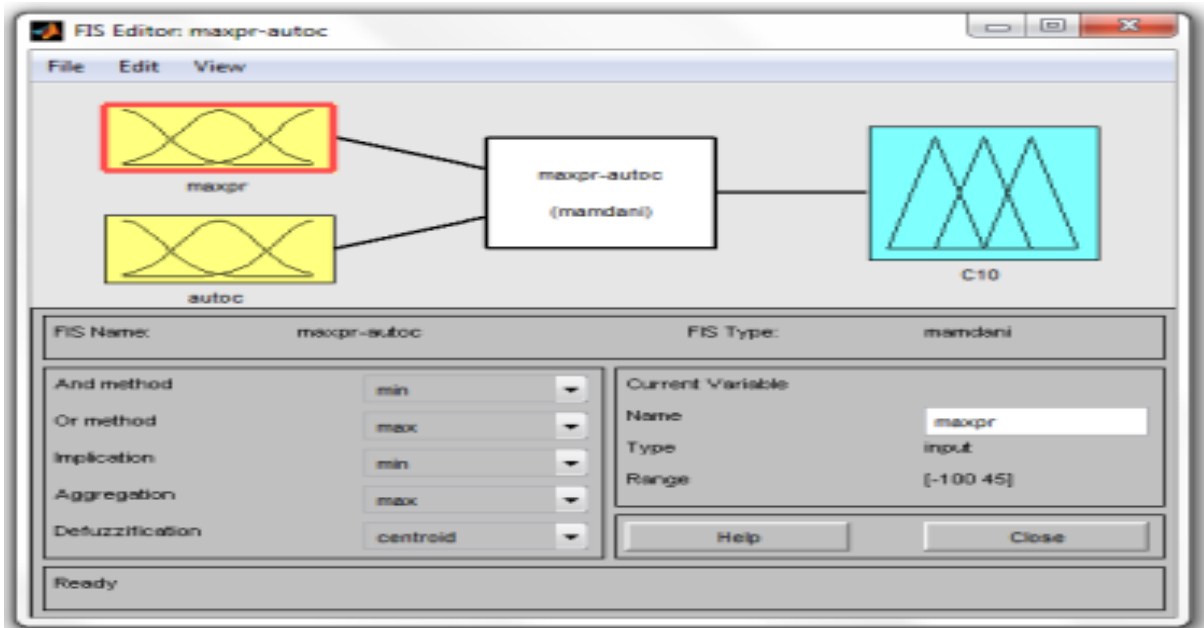


Fig 7: FIS Editor – two inputs

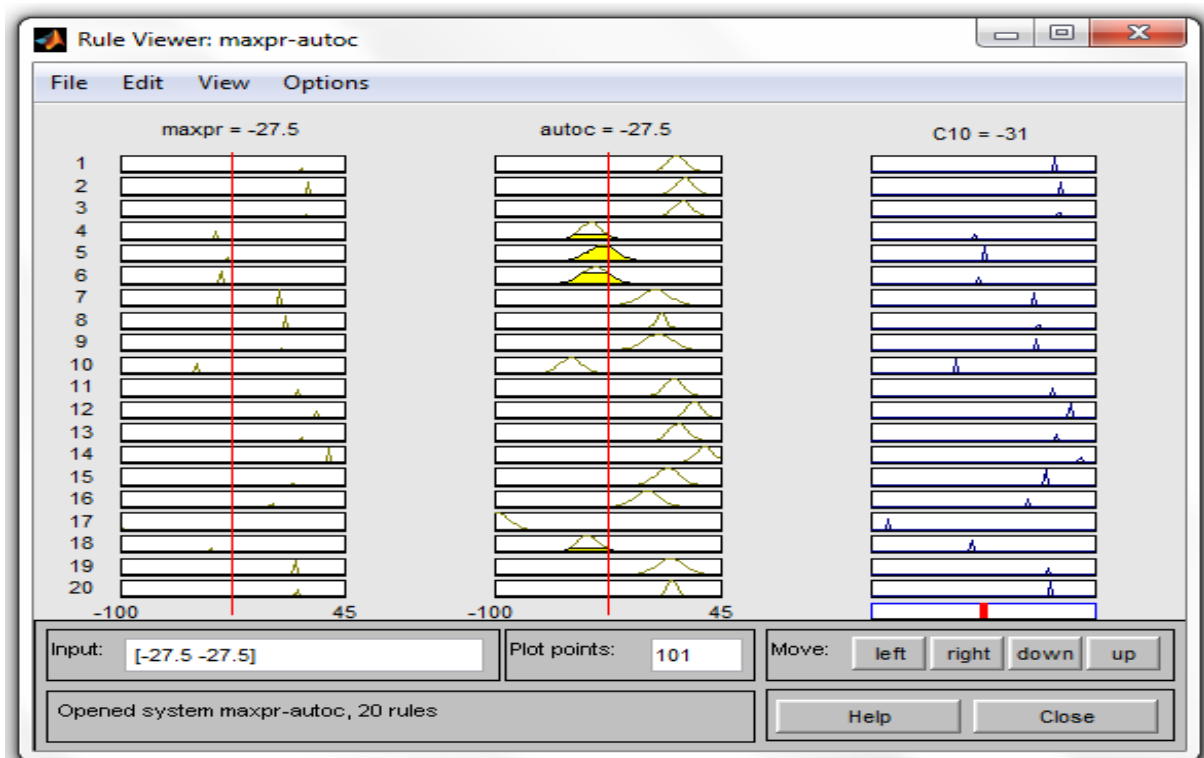


Fig-8: FIS Editor – Rule viewer for two inputs

Where  $x_i$  and  $x_j$  are feature variables,  $\mu_{i,k}(x_i)$  and  $\mu_{j,k}(x_j)$  are defined in Eqn. (1),  $S_{i,j}(x_i, x_j)$  a two-dimensional surface in feature space corresponding to a fuzzy rule:

IF  $x_i$  is  $\mu_{i,k}(x_i)$  and  $x_j$  is  $\mu_{j,k}(x_j)$ , THEN  $y$  is  $y_k$ .

Unlike with fuzzy curves, the surface values calculated by using formulae given in Eqn. (4) for the fuzzy surfaces are used to evaluate the approximation of the input output data.



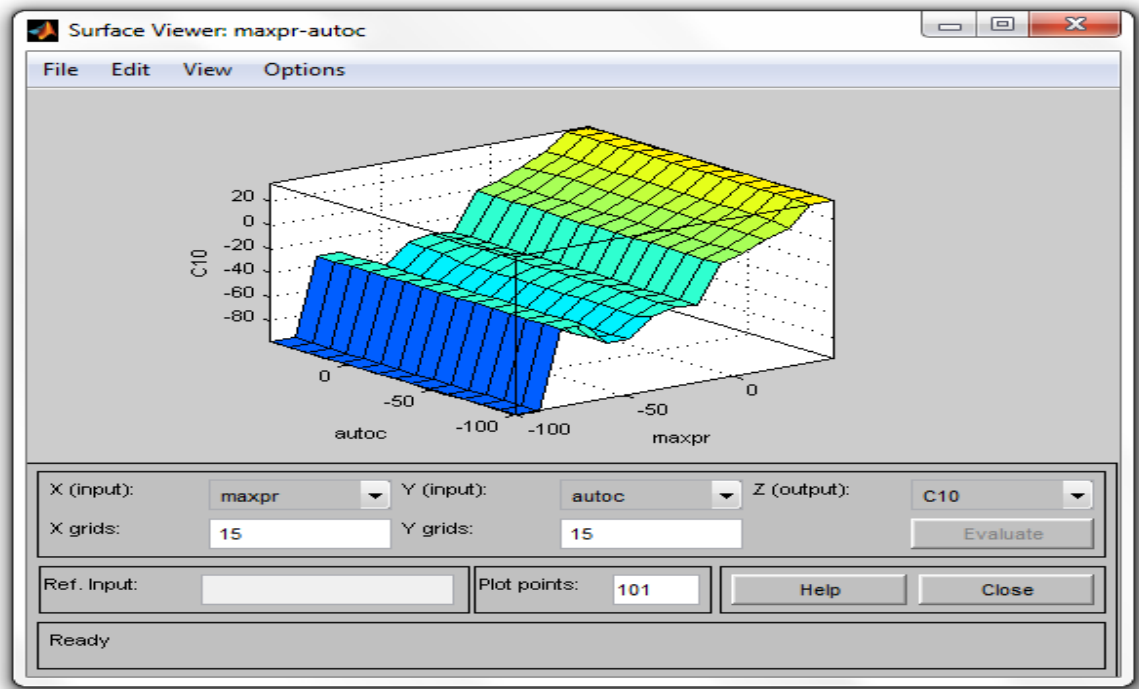


Fig-9: FIS Editor – Surface viewer for two inputs

Now use the most important feature and each of other remaining features, chosen by the fuzzy curve, to create fuzzy surfaces, rank the features in order of dependence with the most important feature (in descending surface value) [35]. Use the second most important feature and each of remained features in previous step to create new fuzzy surfaces, once again. Order the features which are used to create fuzzy surface with the second most important feature by the surface value of their fuzzy surfaces. The third important feature is identified and another 20% of the input features with largest mean square errors are eliminated. Repeat fuzzy surfaces process until enough important inputs are obtained, or no remaining inputs are left.

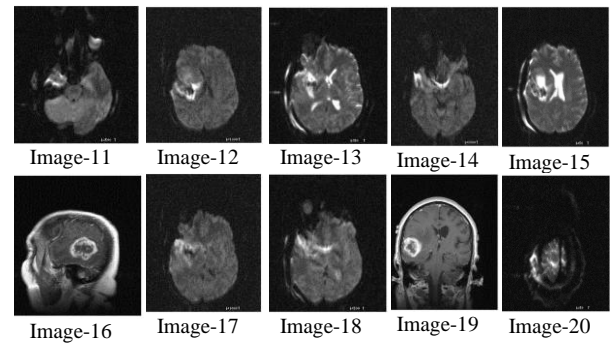


Fig-10: Original Brain MR images

#### 4. RESULTS AND DISCUSSIONS

Results obtained by features selection algorithms are shown in Fig 10 and Fig 11. Database used for this work has been taken from the Rockland Hospital, B-33, 34, Tara Crescent Road, Qutab Institutional Area, New Delhi, Delhi - 416887, India, American National Cancer Data Base (NCDB) [36] and National Cancer Institutes' Comprehensive Cancer Database [37]. Twenty brain MR images have been used in this work.

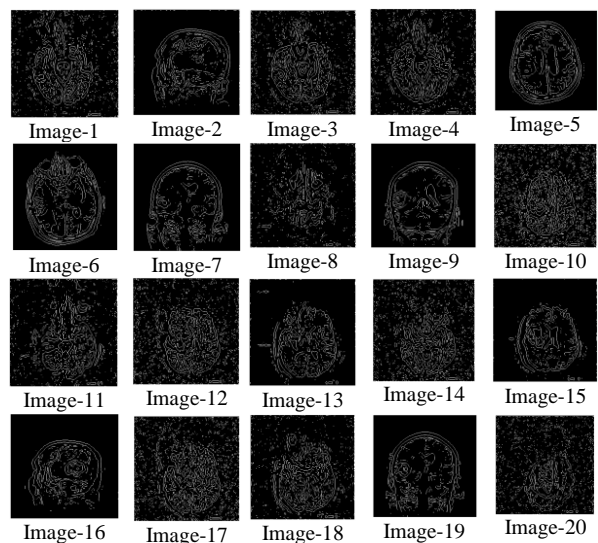
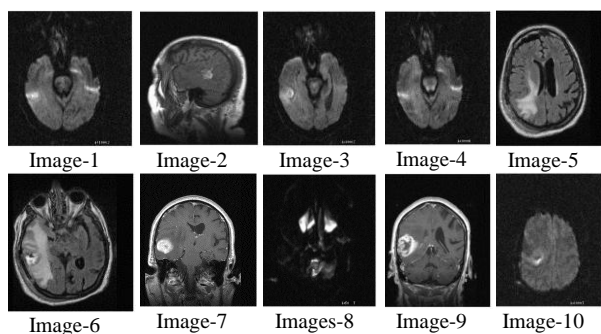
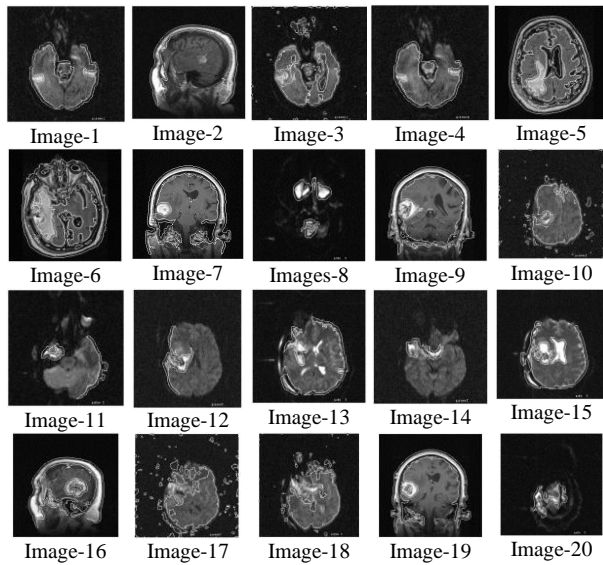


Fig-11: Filtered Brain MR images



**Fig-12: Segmented Brain MR images**

By using GLCM we extracted textural features of each image which are shown in Tables 3-6 below.

**Table 3: Extracted Features of images**

Features	Images					
	I-1	I-2	I-3	I-4	I-5	I-6
autoc	5.73	5.54	5.53	6.55	8.73	8.58
contr	1.48	1.32	1.71	2.09	2.71	2.91
corrmm	0.63	0.64	0.56	0.61	0.62	0.59
cprom	221.8	204.4	177.0	412.0	390.9	359.2
cshad	25.70	24.21	21.48	42.27	36.54	33.66
dissi	0.56	0.54	0.59	0.69	0.8	0.88
energ	0.16	0.16	0.16	0.15	0.15	0.14
entro	2.19	2.18	2.21	2.28	2.62	2.74
homom	0.78	0.78	0.78	0.76	0.74	0.73
maxpr	0.26	0.27	0.26	0.27	0.33	0.33
sosvh	6.40	6.14	6.33	7.52	10.01	9.96
savgh	4.22	4.15	4.20	4.42	5.09	5.08
svarh	12.24	11.65	11.74	15.09	20.80	20.10
senth	1.83	1.82	1.84	1.90	2.07	2.13
denth	0.91	0.90	0.92	1.02	1.19	1.25
inf1h	-0.33	-0.32	-0.32	-0.31	-0.29	-0.27
inf2h	0.76	0.75	0.75	0.75	0.77	0.75
indnc	0.94	0.94	0.94	0.93	0.92	0.91
idmnc	0.98	0.98	0.98	0.97	0.96	0.96

**Table 4: Extracted Features of images**

Features	Images					
	I-7	I-8	I-9	I-10	I-11	I-12
autoc	9.06	3.10	9.16	6.69	5.90	4.74
contr	2.21	1.00	2.03	1.71	0.84	1.00
corrmm	0.68	0.60	0.70	0.63	0.72	0.67
cprom	450.8	366.0	484.7	357.9	152.9	156.5
cshad	39.80	33.42	42.60	35.33	17.64	20.28
dissi	0.71	0.47	0.68	0.55	0.43	0.50
energ	0.14	0.30	0.14	0.17	0.17	0.18
entro	2.59	1.60	2.63	2.15	2.13	2.06
homom	0.76	0.80	0.77	0.81	0.81	0.78

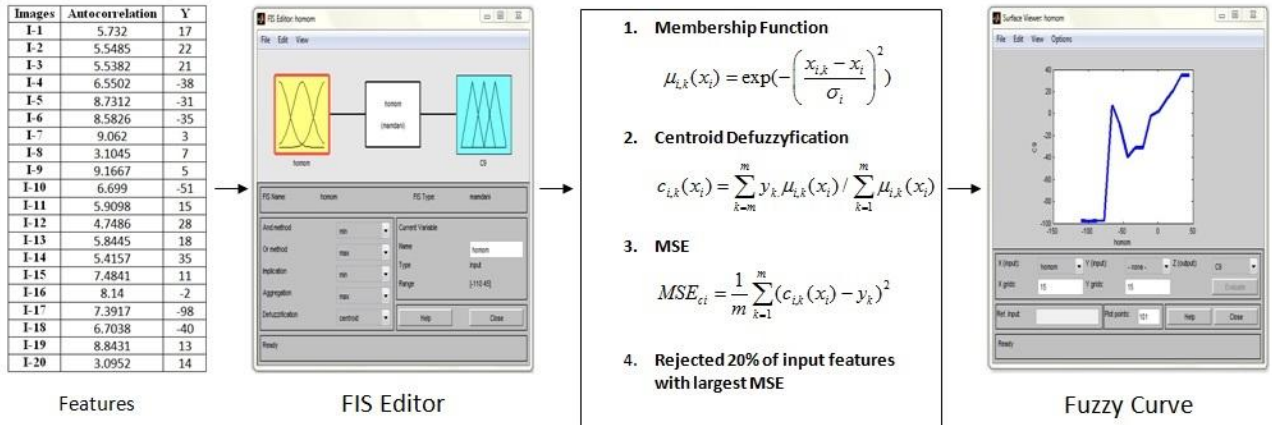
maxpr	0.30	0.49	0.30	0.28	0.31	0.31
sosvh	10.08	3.56	10.10	7.48	6.26	5.18
savgh	5.17	3.04	5.18	4.55	4.37	3.84
svarh	21.03	7.19	21.1	15.0	11.84	9.72
senth	2.11	1.29	2.13	1.86	1.82	1.71
denth	1.10	0.81	1.09	0.89	0.82	0.84
inf1h	-0.32	-0.16	-0.32	-0.38	-0.37	-0.31
inf2h	0.79	0.49	0.80	0.79	0.79	0.73
indnc	0.93	0.95	0.93	0.94	0.95	0.94
idmnc	0.97	0.98	0.97	0.98	0.98	0.98

**Table 5: Extracted Features of images**

Features	Images					
	I-13	I-14	I-15	I-16	I-17	I-18
autoc	5.84	5.41	7.48	8.14	7.39	6.70
contr	1.35	0.65	1.39	1.87	2.57	1.8
corrmm	0.61	0.75	0.74	0.73	0.54	0.61
cprom	156.7	145.1	494.1	634.5	387.8	348.3
cshad	18.36	16.57	46.94	60.74	37.44	34.68
dissi	0.48	0.41	0.52	0.67	0.70	0.57
energ	0.17	0.19	0.16	0.16	0.14	0.17
entro	2.07	2.06	2.32	2.52	2.34	2.15
homom	0.82	0.81	0.81	0.76	0.78	0.80
maxpr	0.29	0.33	0.27	0.35	0.27	0.28
sosvh	6.45	5.67	8.11	9.00	8.60	7.56
savgh	4.35	4.20	4.66	4.72	4.84	4.57
svarh	12.35	10.61	16.76	19.23	16.76	15.09
senth	1.78	1.76	1.98	2.05	2.00	1.86
denth	0.82	0.79	0.92	1.06	1.00	0.90
inf1h	-0.39	-0.35	-0.39	-0.32	-0.34	-0.38
inf2h	0.79	0.76	0.82	0.79	0.78	0.79
indnc	0.95	0.95	0.94	0.93	0.93	0.94
idmnc	0.98	0.99	0.98	0.97	0.97	0.97

**Table 6: Extracted Features of images**

Features	Images	
	I-19	I-20
autoc	8.84	3.09
contr	2.00	0.94
corrmm	0.69	0.54
cprom	432.4	179.0
cshad	38.52	19.44
dissi	0.67	0.50
energ	0.14	0.24
entro	2.58	1.71
homom	0.77	0.78
maxpr	0.30	0.38
sosvh	9.76	3.51
savgh	5.11	3.17
svarh	20.27	6.44
senth	2.10	1.38
denth	1.07	0.83
inf1h	-0.32	-0.17
inf2h	0.79	0.5
indnc	0.93	0.94
idmnc	0.97	0.98



### FUZZY CURVES

Fig-13: Features section procedure using fuzzy curves

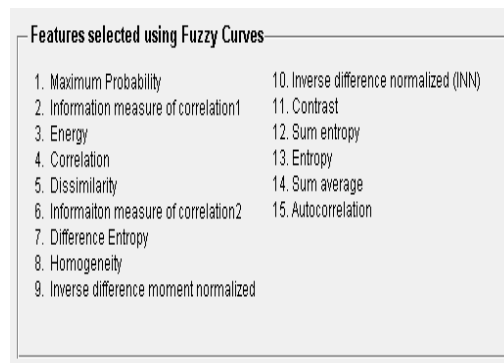
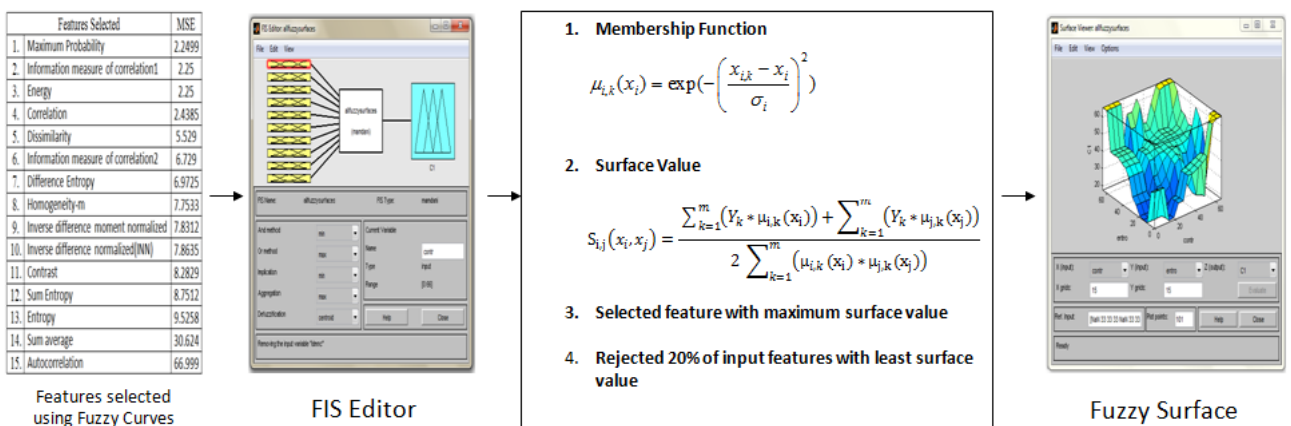
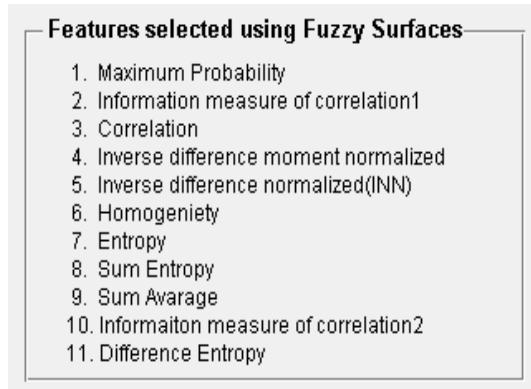


Fig-14: Selected features using fuzzy curves



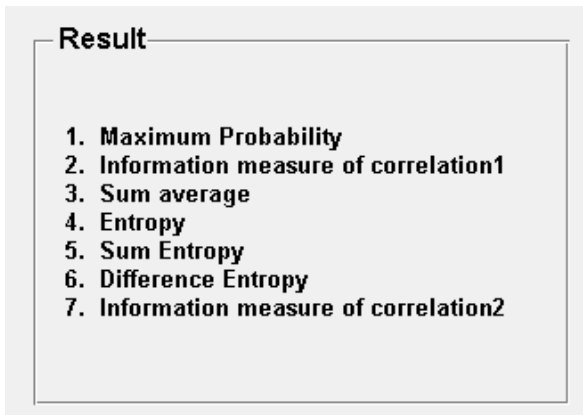
### FUZZY SURFACES

Fig-15: Features selection procedure using fuzzy surfaces



**Fig-16: Selected features using fuzzy surfaces once**

By repeating fuzzy surfaces process seven important features are selected from the extracted set of nineteen textural features.



**Fig-17: Finally selected features using fuzzy surfaces thrice**

## 5. CONCLUSIONS

This work proposes for the first time a new fuzzy features selection technique for brain MR images. Features are extracted using GLCM and images are pre-processed before segmentation. These features have been selected using fuzzy curves which isolate the significant features from the complete set of features extracted using GLCM and fuzzy surfaces which are based on the principle that independent features do better work rather than dependent features. The relevance of fuzzy subset selection method relies on the fact that it takes into consideration the particular characteristics of the fuzzy logic used to generate the interpretation of domain such as the fuzzy knowledge base. The approach is tested on twenty brain MR images and results show that it is able to achieve a good reduction in the number of features as well as keeping very low error rates comparable to existing feature selection methods.

Fuzzy feature selection approach can be further improved in future by combining genetic algorithm approach with it.

## 6. REFERENCES

- [1] Brain Tumor Facts – ABTA, available at [www.abta.org/news/brain-tumor-fact-sheets](http://www.abta.org/news/brain-tumor-fact-sheets)
- [2] F. C. Morabito, “A fuzzy neural approach to localizing holes in conducting plate”, IEEE Transactions on Magnetics, Vol. 37, No. 5, 2001.
- [3] H. Narazaki, and A. L. Ralescu, “An improved synthesis method for multilayered neural networks using qualitative knowledge,” IEEE Trans. Fuzzy System., Vol. 1, No. 2, pp. 125–137, 1993.
- [4] SPP11 statistics analysis tutorial. Beijing Hope Electronics Press, Beijing, 2002.
- [5] S. Winkler, M. Affenzeller, and S. Wagner, “Genetic programming based model structure identification using on-line system data”, Conceptual Modeling and Simulation Conference – CMS, , Marseilles, France Oct. 2005.
- [6] Yinghua Lin, and G. Cunningham, “A fuzzy approach to input variable identification”, Proceeding of 3rd IEEE International Conference on Fuzzy Systems, Orlando, pp. 2031-2036, Jun.1994.
- [7] S.Javeed Hussain, T.Satya Savithri, P.V Sree Devi, “Segmentation of tissues in brain MR images using dynamic neuro-fuzzy technique”, IJSCE,vol.-I, issue-6, January 2012.
- [8] Indah Soesanti, Adhi Susanto, Thomas Sri Widodo, Maesadji Tjokronagoro, “Optimized fuzzy logic application for MRI brain images segmentation”, International Journal of Computer Science and Information Technology (IJCSIT) vol. 3, No. 5, Oct 2011.
- [9] Parmida Moradi Birgani, Meghdad Ashtiyani, Saeed Asadi, “MRI segmentation using fuzzy C-means clustering algorithm basis neural network”, Tehran, Iran. ] Information and Communication Technologies: From Theory to Applications, 2008. ICTTA 2008.
- [10] ZHANG Nan, “Feature selection based segmentation of multi-Source images: application to brain tumor segmentation in multi-sequence MRI”, © [N. Zhang], [2011], INSA de Lyon, tous droits reserves
- [11] J. Jiang, P. Trundle\*, J. Ren, ”Medical image analysis with artificial neural networks”, Computerized Medical Imaging and Graphics 34 (2010) 617–631.
- [12] D. Dumitrescu, B. Lazzerini, and L. Jair, “Fuzzy sets and their application to clustering and training”, ser. international series on computational intelligence. CBC Press, 2000.
- [13] W. Pedrycz, Fuzzy modelling: paradigms and practice. kluwer academic press, 1996.

- [14] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modeling," *IEEE Trans. Fuzzy Systems*, vol. 1, no. 1, 1993, 7-31.
- [15] H. Takagi, I. Hayashi, "-driven fuzzy reasoning," *Int 'l. J. approximator*, vol. 5, no. 3, 1991, 191-212.
- [16] J.B. Antoine Maintz, Petra A. van den Elsen and Max A. Viergever, "Registration of SPECT and MR brain images using a fuzzy surface", Radiological Science Laboratory, Stanford University School of Medicine, Stanford, CA, USA.
- [17] Maria-Elena Algorri and Fernando Flores-Mangas, "Classification of anatomical structures in MR brain images using fuzzy parameters", *IEEE trans. VOL-51*, Sept. 2004.
- [18] M. E. Cintra, H. A. Camargo, and M. C. Monard, "A study on techniques for the automatic generation of membership functions for pattern recognition," *C3N - Congresso da Academia Trinacional de Ciências*, vol. 1, pp. 1–10, 2008.
- [19] T. W. Liao, A. K. Celmins, and R. J. Hammell, "A fuzzy C-Means variant for the generation of fuzzy term sets," *Fuzzy Sets and Fuzzy Systems*, vol. 135, pp. 241–257, 1997.
- [20] S. R. Jang, C. T. Sun, and E. Mizutani, *Neuro-fuzzy and Soft computing*. Prentice Hall, 1997.
- [21] O. Cordón, F. A. C. Gomide, F. Herrera, F. Hoffmann, and L. Magdalena, "Special issue on genetic fuzzy systems," *Fuzzy Sets and Systems*, vol. 141, 2004.
- [22] H. Ishibuchi and Y. Nojima, "Evolutionary multiobjective optimization for the design of fuzzy rule-based ensemble classifiers", *International Journal of Hybrid Intelligent Systems*, vol. 3, pp. 129–145, 2006.
- [23] O. Cordón, R. Alcalá, J. Alcalá-Fdez, and I. Rojas, "Special section on genetic fuzzy systems," *IEEE Transactions on Fuzzy Systems*, vol. 15, pp. 533–592, 2007.
- [24] F. Hoffmann, "Combining boosting and evolutionary algorithms for learning of fuzzy classification rules," *Fuzzy Sets and Systems*, vol. 141, pp. 47–58, 2004.
- [25] P. A. D. Castro and H. A. Camargo, "Focusing on interpretability and accuracy of genetic fuzzy systems," *IEEE International Conference on Fuzzy Systems*, vol. 1, pp. 696–701, 2005.
- [26] M. E. Cintra, H. de Arruda Camargo, and M. C. Monard, "Fuzzy feature subset selection using the Wang & Mendel method", *Hybrid Intelligent Systems - HIS 2008*, vol. 1, 2008.
- [27] L. Wang, "The WM method completed: a flexible fuzzy system approach to data mining," *IEEE International Conference on Fuzzy Systems*, vol. 11, pp. 768–782, 2003.
- [28] H.P. Chan, K. Doi, and C.J. Vyborny. "Improvements in radiologists' detection of clustered micro-calcifications on mammograms: the potential of computer aided diagnosis". *Investigative Radiology*, 25:1102{1110, 1990.
- [29] David A. Clausi, "An analysis of co-occurrence texture statistics as a function of gray level quantization", *Can. J. Remote Sensing*, Vol. 28, No. 1, pp. 45-62, 2002.
- [30] C.Callins Christiyana, V.Rajamani and A.Usha devi, "Ultra sound kidney image retrieval using time efficient one dimensional GLCM texture feature" , Special Issue of International Journal of Computer Applications (0975 – 8887) on Advanced Computing and Communication Technologies for HPC Applications - ACCTHPCA, June 2012.
- [31] David A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization", *Can. J. Remote Sensing*, Vol. 28, No. 1, pp. 45–62, 2002.
- [32] Yinghua. Lin, G. Cunningham, S. V. Coggeshall, and R. D. Jone, "Nonlinear system input structure identification: two stage fuzzy curves and surfaces", *IEEE Transactions on System, Man, and Cybernetics—Part A: System and Humans*, Vol. 28, No. 5, pp. 678-684, Sept. 1998.
- [33] Yan-cheng zhang, Han-ping mao, Bo hu, Ming-xi li, "Features Selection Of Cotton Disease Leaves Image Based On Fuzzy Feature Selection Techniques", *Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition*, Beijing, China, 2-4 Nov. 2007-4244-1066-5.
- [34] Panagiotis. T., Stelios E. P., and Dimitris M., "Plant leaves classification based on morphological features and a fuzzy surface selection technique", *proceeding of 5th ICTA, Thessaloniki, Greece*, pp. 365-370, Oct. 2005.
- [35] Bin Li, "Fuzzy approach to identification of input variable", *Journal of Harbin Engineering University*, Vol. 24 No. 1, pp. 104-107, Feb. 2003 (in Chinese).
- [36] National Cancer Data Base (NCDB) - American College of Surgeons, available at <http://www.facs.org/cancer/ncdb/>
- [37] PDQ® - NCI's Comprehensive Cancer Database - National Cancer Institute, available at: <http://www.cancer.gov/cancertopics/pdq/cancerdatabase>