Genetic Algorithm Based Brain Tumor Detection and Segmentation

K.VishnuLakshmi", PG Scholar, Department of Medical Electronics, Sengunthar College of Engineering, Tiruchengode

Muhamandu Sathik Raja", Assistant Professor, Department of Medical Electronics, Sengunthar College of Engineering, Tiruchengode

Abstract- Tumor is swelling of the body part, generally without any inflammation that happens due to abnormal growth of cells in that place of the body. Brain tumor is difficult to diagnose at initial stage. The tumor is diagnosed by magnetic resonance imaging (MRI) and depending on it; the tumors are distinguished into different grades of severity. In early a neuroanist based tumor detection approach is proposed to tumor diagnosing at initial stage but it is not give better solution. So we propose a new method to detect and extract tumor from the whole images. This paper proposes a scheme for extracting tumor from MRI based on GA (Genetic Algorithm) based FCM (Fuzzy C-means) clustering and morphological operation. Computer simulations of our algorithm present a considerable improvement over other existing techniques.

Keywords: FCM, Genetic Algorithm, Image Segmentation, Morphological operation

INTRODUCTION

Most of the tumor is two types namely benign and malignant. Malignant tumor is referred to as cancer. Abnormal growth of cell inside brain is called brain tumor. There are two general groups of brain tumor. Primary brain tumor starts in brain and tends to stay there. Secondary brain tumor starts somewhere else in the body but travels to brain. Secondary tumors are more common than primary tumors. The reason for brain tumor is unknown till now. It is considered that probable reasons of brain tumor can be a number of conditions like neurofibromatosis, exposure to chemical vinyl chloride, Epstein-Barr virus and ionizing radiation. The use of mobile phones is also considered as one of the risk factors but there is still no clear evidence. Meningioma (usually benign), Oligodendrogliomas and astrocytoma such as Glioblastomas are primary tumor commonly found in adults and Medulloblastoma in children. Diagnosis is usually done by medical Examination along with MRI. Biopsy is then conducted for confirmation. Tumours are divided into different grades of severity depending on the report obtained from diagnosis. In grade 1, the cells look normal and grow slowly. The cells look slightly abnormal and grow slowly in grade 2. Cell starts to grow actively and look abnormal. They start to grow actively in nearby brain tissue and tend to reoccur. This happens only when it is grade 3 tumor. In grade 4 tumor are most abnormal and fast spreading. There are various literatures available on the topic of brain tumor detection and extraction on MRI images of brain. In recently, a two methods for segmentation, i.e., ACO hybrid with Fuzzy and HSOM hybrid with fuzzy is used to detect brain tumor. Though the detection is done, still the noise is remaining in the image. And another method named as ridier’s method is used to low pass filtering morphological segmentation operation and thresholding and finally region growing methods to extract the brain tumor. But the size of the tumor is not accurate as that were present in original pictures. And also another method a multi-modality framework and AdaBoost classifier is used to detect the tumor. Even if the tumor is detected, it still has noise in it and the accuracy of the detected tumor is not good. It is not like that present in the ground image.

So we use use dynamic GA based image segmentation and Morphological operation is used to detect and extract tumor from MRI images. This will help in easy detection of tumor and also see the growth and size of the tumor. First, the FCM based GA technique is used to segment the image into F number of clusters. The FCM is used because it has better noise removal power. Then the thresholding is applied on segmented images and the region containing the tumor is clearly detected. Finally, morphological operation is used so that only the tumor can be extracted from the whole image while keeping the position, the size and the area intact.

II. EXISTING METHOD

2.1. ANFIS BASED TUMOR CLASSIFICATION

Adaptive Neuro Fuzzy Inference Systems (ANFIS) are adaptive networks implemented in MATLAB software. An adaptive network consists of a group of nodes and directional links.
ANFIS is a kind of neural network that is based on Takagi–Sugeno fuzzy inference system. Since it integrates both neural networks and fuzzy logic principles, it has potential to capture the benefits of both in a single framework. Its inference system corresponds to a set of fuzzy IF–THEN rules that have learning capability to approximate nonlinear functions. Hence, ANFIS is considered to be a universal estimator.

Fuzzy systems present particular problems to a developer:

• Rules: The if-then rules have to be determined somehow. This is usually done by ‘knowledge acquisition’ from an expert. It is a time consuming process that is fraught with problems.

• Membership functions: A fuzzy set is fully determined by its membership function. This has to be determined. The network has a learning rule such as back propagation. ANFIS is said to be adaptive because some or all the nodes have parameters which influences the output node. ANFIS is a supervised learning technique and relates the inputs with the outputs.

Consider an ANFIS classifier based on a Two-rule system as below,

\[ \text{If } x \text{ is } A1 \text{ and } y \text{ is } B1 \text{ then, } f1 = px + qy + r1 \]

\[ \text{If } x \text{ is } A2 \text{ and } y \text{ is } B2 \text{ then, } f2 = px + qy + r2 \]

For the training of the network, there is a forward pass and a backward pass. We now look at each layer in turn for the forward pass. The forward pass propagates the input vector through the network layer by layer. In the backward pass, the error is sent back through the network in a similar manner to back propagation.

2.2. LAYERS OF ANFIS ARCHITECTURE

LAYER 1

At layer 1, the output of each node is given by,

\[ o_{1,i} = \mu_{A_i}(x) \text{ for } i = 1, 2 \]

\[ o_{1,i} = \mu_{B_{i-2}}(y) \text{ for } i = 3, 4 \]

Where, \( o_{1,i}(x) \) is essentially the member for \( x \) and \( y \). The membership function we have considered is a bell shaped function given by:

\[ \mu_{A_i}(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}} \]

(a.1)
Where \(a_i, b_i, c_i\) are the ground parameters to be trained.

**Layer 2**
In this layer, all the nodes are fixed. Here, the t-norm is used to ‘AND’ the membership grades, for example, the product given by,

\[ o_{2,i} = w_i = \mu_{A_i}(x) \mu_{B_i}(y), \quad i = 1, 2 \]  

(a.2)

**Layer 3**
Layer 3 contains fixed nodes which calculate the ratio of the firing strengths of the rules, as given below,

\[ o_{3,i} = \frac{w_i}{w_1 + w_2} \]  

(a.3)

**Layer 4**
In this layer, the nodes are adaptive and achieve the consequent of the rules:

\[ o_{4,i} = \frac{w_i f_i}{w} = w_i \left( p_i x + q_i y + r_i \right) \]  

(a.4)

The parameters in this layer \((p_i, q_i, r_i)\) are to be determined and are referred to as the consequent parameters.

**Layer 5**
In this layer, a single node is present which computes the overall output.

\[ o_{5,i} = \sum_i w_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \]  

(a.5)

Then, the input vector is fed through the network layer by layer. We now consider how the ANFIS learns the premise and consequent parameters for the membership functions and the rules.

**2.3. Operation of ANFIS Classification**
For an ANFIS adaptive network with fixed premise parameters, the output is linear in the consequent parameters. The total parameter can be categorized into three:

- \(S\) = set of total parameters
- \(S_i\) = set of premise parameters
- \(S_o\) = set of consequent parameters

ANFIS employs a 2-pass supervised learning algorithm, which consists of, a Forward Pass and a Backward Pass. The steepest descent algorithm is used along with the least squares algorithm to adapt to the parameters in this ANFIS algorithm.

**2.3.1. The Forward Pass**
In the forward pass, neuron outputs are calculated layer by layer and the consequent parameters are identified by the least squares estimator (LSE) to obtain the final single output. The forward pass operation on each layer is described below: Layer 1 known as a fuzzification layer, define the membership grades for each set of input and depends on the fuzzy membership function chosen. Here \(S_i\) is fixed and \(S_o\) is estimated using a Least Squares algorithm. The procedure is as follows:

- Step 1: Feed the input vector
- Step 2: Compute the output at each node and at each layer
- Step 3: Repeat for all data until y is produced
- Step 4: Find the parameters in \(S_o\) using Least Squares
- Step 5: Calculate the error for each training pair

**2.3.2. The Backward Pass**
In the backward pass, error signals are propagated and the antecedent parameters are updated according to the chain rule. According to the gradient descent (GD), the error rate for training data is calculated. Here \(S_o\) is fixed and \(S_i\) is estimated using back propagation. The parameters in \(S_i\) are updated by back propagation for given fixed values of \(S_o\) and these parameters are assured to be the global optimal point. ANFIS is a model machine learning methodology, which derives its behaviour from a metaphor of the processes of evolution in nature and are able to overcome complex non-linear optimization tasks like non-convex problems, non-continuous objective functions, etc. It is based on an initial random population of solutions and an iterative procedure, which improves the characteristics of the population and produces solutions that are closer to the global optimum. This is achieved by applying a number of genetic operators to the population, in order to produce the next generation of solutions. Genetic Algorithm may be used successfully in combinations with neural and fuzzy systems.
Particularly in neurofuzzy control, GAs have been utilized extensively to tune the neurofuzzy controller parameters and acquire the fuzzy rules.

The structure of adaptive network can also be reduced into four layers, where the output of each rule is induced jointly by the output membership function and the firing strength. This structure is known as type-1 ANFIS. ANFIS training can use alternative algorithms to reduce the error of the training. A combination of the gradient descent algorithm and a least squares algorithm is one of the methods for an effective search for the optimal parameters and to further speed up the learning process. Jang & Mizutani proposed the Levenberg-Marquardt (LM) method for ANFIS learning and compared this experimentally to the previously proposed hybrid ANFIS. Although the LM method gave a better RMSE after the predetermined iterations, the final membership functions varied a lot from the initial setting consequently it is hard to assign appropriate linguistic interpretations.

There are other algorithms that have been implemented to update the antecedent of the ANFIS network, such as, gradient descent (GD), resilient propagation (RPROP), Quickprop (QP) and LM to update the membership functions, while the consequent part remains as the recursive least square estimation. The Fletcher-Reeves, scaled conjugate gradient (SCG) and particle swarm optimization have also been used for tuning parameters of ANFIS membership functions.

### III. PROPOSED METHOD

#### 3.1 GENETIC ALGORITHM BASED TUMOR CLASSIFICATION

The aim of this paper is to detect and extract tumors from different types of brain tumor images. Therefore, the following methods have been used to cater to the purpose:

- **FCM based GA**
- **Thresholding**
- **Morphological operations**

#### 3.2. FCM BASED GENETIC ALGORITHM

The searching capacity of GA is used to appropriately cluster a set of n unlabeled points in N dimensions into F clusters. The GA will help the points found out by the FCM clustering method to converge into a single set of values. The steps for performing the algorithm are given as follows.

##### 3.2.1. ENCODING

Every chromosome consists of a sequence of F cluster centers. Each cluster center is mapped to G consecutive genes in the chromosome in G dimensional space. For image datasets each gene is an integer representing an intensity value.

##### 3.2.2. POPULATION INITIALIZATION

Genetic algorithm needs a population of size S to work on. This population is generated using the FCM clustering method. The FCM method is executed S times to get the population of size S. Each chromosome obtained in S represents the no. of clusters c=F.

The membership values of the data points with each cluster center are calculated and clustering operation is carried out. The membership values assigned to the data centers are done on the basis of distance between the data points and its cluster centers. A data point near to a cluster center signifies the membership of that data point to its corresponding center is more. Let \( X = \{x_1, x_2, x_3, \ldots, x_n\} \) be the set of data points and \( UC = \{u_1, u_2, u_3, \ldots, u_c\} \) be the set of centers. Randomly ‘c’ clusters centers are selected.

The fuzzy membership value ‘\( \eta_{ij} \)’ using:
\[ n_{ij} = \frac{1}{\sum_{k=1}^{F} \left( \frac{n_{ik}}{\mu_k} \right)^\alpha} \quad (1) \]

where,

\[ e_{ij} = \| x_i \Box y_j \| \quad \text{and} \quad e_{ik} = \| x_i \Box y_k \| \]

Compute the fuzzy centers ‘UCj’ using:

\[ UC_j = \frac{\sum_{i=1}^{n} (n_{ij})^\alpha}{\sum_{i=1}^{n} n_{ij}}, \quad \text{for all } j = 1,2,\ldots, F \quad (2) \]

The points start to converge gradually while implementing the algorithm

### 3.2.3. Fitness Computation

The calculation of fitness consists of two phases. In first phase clustering of data points is done according to the cluster centers which will become chromosome set. The intensity value \( x_i, i=1,2,\ldots,m \times n \) is assigned to cluster with center \( y_j, j=1,2,\ldots,F \).

\[ \text{if } \| x_i - y_j \| < \| x_i - y_p \|, \quad p = 1,2,\ldots,F \quad \text{and} \quad p \neq F \quad (3) \]

In the next phase, the values of the cluster centers are updated the mean points of respective clusters which are encoded as chromosome.

\[ \bar{y}^t_i = \frac{1}{n_i} \sum_{x_{jt} \in S_i} x_{jt} \quad t = 1,2,\ldots,F \quad (4) \]

Where \( \bar{y}_i \) is new center for cluster \( C_i \).

\( S \) is the summation of Euclidean distance of each point from their respective clusters is given by

\[ s = \sum_{i=1}^{N} s_i \quad (5) \]

\[ S_t = \sum_{x_{jt} \in S_i} \| x_{jt} - \bar{y}_i \| \quad (6) \]

The Fitness function is given by:

\[ \omega = \frac{1}{s} \quad (7) \]

### 3.2.4. Selection

The selection of the best chromosomes in the population is used by Roulette Wheel selection; it is proportional to the fitness of that chromosome. This means the chromosome having the highest fitness has the maximum probability to get selected for the consequent processes

\[ \rho_l = \frac{\omega_{l}}{\sum_{l=1}^{N} \omega_l} \quad (8) \]

### 3.2.5. Crossover

The chromosomes are selected through Roulette Wheel selection process, for better improvement of the fitness value the chromosomes perform crossover operation. On the selected pair of chromosomes single point crossover is applied. On the selected pair, a point from the whole length of chromosome is chosen randomly and the whole content beyond that point is interchanged to get the offspring.

### 3.2.6. Mutation

Chromosomes obtained after crossover undergoes mutation which has a fixed probability \( \mu_m \). Here a point in the pair of chromosome is selected randomly and the content of that point is interchanged. The pair will be chosen after the crossover process is completed.

### 3.2.7. Termination Criterion

The above mentioned four steps, fitness function calculation, roulette wheel selection, crossover and mutation is repeated until it encounters a predefined number of steps. The result of each iteration is updated and saved. At the end of the whole process a chromosome is obtained which will give the best result.

### 3.3. Thresholding

Finally, the best chromosomes set that is obtained is used as the centers points of the cluster, and get final segmented image. After that the thresholding is applied on the segmented image to generate a binary image having two values 0 and 255.
A value $\gamma$ is chosen and every pixel that has intensity value less than $\gamma$ is made 0 otherwise 255. This operation helps in identifying the region that contains tumor and also helps in extracting it. Hence a binary image $I$ is constructed, where $g$ is original image and $\gamma$ is threshold value.

$$I(x) = \begin{cases} 0 & \text{if } g(x) \leq \gamma \\ 1 & \text{if } g(x) > \gamma \end{cases}$$

### 3.4. MORPHOLOGICAL OPERATION

Any shape and morphology feature related operation of images, mostly morphological operation is used. In this paper, the unwanted non-tumor portions of the MRI images are removing by morphological operations. For this purpose, we apply morphological opening operation. Dilation and erosion are combined together to form the opening operations. The dilation, erosion and opening operation are given by:

- **Dilation:** $D(f, g) = f \oplus g = \bigcup_{a \in g} (f - a)$
- **Erosion:** $E(f, g) = f \ominus g = \bigcap_{a \in g} (f - a)$
- **Opening:** $O(f, g) = f \circ g = D(E(f, g) \ominus g)$

After performing these steps two other steps are conducted. They are to determine the connected components then computing area of each component and then removing the small objects. The overview of the method is given in Fig.1

![Block diagram of the proposed method](image)

**IV. EXPERIMENT RESULT**

Extensive experiments are performed on a variety of standard tumor affected T2 weighted MRI images with distinctly different features and different sizes. The outcome of the different MRI images using proposed method is shown in Fig.2. Fig.2 (b), (e), (i), (l), (o) are shown the extracted tumor from original images using our proposed method and Fig.2 (c), (f), (j), (m), (p) are also shown the original image where tumor is detected and it is identified by red color. Our proposed scheme gives better results than K-means and KFCM [6]. The performance of each of the algorithms is evaluated by the number of missed alarms (MA), false alarms (FA) and overall errors (OE). From this result, it is shown that the accuracy of the proposed method is better than the K-means and KMCM (in Fig.3 and Fig.4). Overall error and accuracy is calculated from the given equation:

$$\text{Overall error} = \frac{FA + MA}{\text{Total pixel}}$$

$$\text{Accuracy} = \frac{\text{Total pixel} - \text{Error}}{\text{Total pixel}}$$
V. CONCLUSION

The proposed methodology presents a new method to detect and extract tumor from a MRI of brain images. This proposed method does not need any previous information for segmentation. This detection technique established the advantage and more steady than other tumor detection techniques. In future, this method needs to be improved more for real time applications.

VI. REFERENCES

[7] Sahar Ghanavati, Junning Li, Ting Liu, Paul S. Babyn, Wendy Doda and GeorgeLampropoulos, “Automatic Brain Tumor Detection in Magnetic Resonance Images”, AUG Signals Ltd., Toronto, Canada. 2 Hospital for Sick Children, Department of Medical Imaging, University of Toronto, ON., Canada.