Fuzzy C-Means, ANFIS and Genetic Algorithm for Segmenting Astrocytoma – A Type of Brain Tumor

Minakshi Sharma¹
AP, IT Department
JMIT, Radaur, India

Dr. Sourabh Mukherjee ²
AP, CSE Department
Banasthali University Rajasthan, India

Abstract - Imaging plays an important role in medical field like medical diagnosis, treatment planning and patient follow up. Image segmentation is the backbone process to accomplish these tasks by dividing an image in to meaningful parts which share similar properties. Medical Resonance Imaging (MRI) is a primary diagnostic technique to do image segmentation. There are several techniques proposed for image segmentation of different parts of body like Region growing, Thresholding, Clustering methods and Soft computing techniques (Fuzzy Logic, Neural Network, Genetic Algorithm). The proposed research work uses Grey level Co-occurrence Matrix (GLCM) for texture feature extraction, ANFIS (Adaptive Network Fuzzy inference System) plus Genetic Algorithm for feature selection and FCM (Fuzzy C-Means) for segmentation of Astrocytoma (Brain Tumor) with all four Grades. The comparative study between FCM, FC plus K-mean, Genetic Algorithm, ANFIS and proposed technique shows improved Accuracy, Sensitivity and Specificity.

Keywords: MRI, ANFIS, FCM, Astrocytoma, Grades I to IV, GLCM, K-means

1. INTRODUCTION

Image Segmentation is a process that groups pixel which has similar attributes. The results can be used in medical field for treatment planning, tumor identification, to analyze tumor growth and computer assisted surgery. This can be done in two ways: Manual Segmentation and Automatic Segmentation. Manual segmentation is a versatile technique but can be affected by fatigue and large number of images. While automated methods provide rapid identification of tumor in comparison to manual segmentation and also results are reproducible [1]. There are various image segmentation algorithms like, Edge Based algorithms, Region based algorithms [2][3], data clustering algorithms[4], Thresholding [5][6]. But, proposed research work focus literature survey related only to the algorithms for brain tumor Segmentation. D. Jude Hemanth [7] presents a which are used comparative study between ANFIS, FCM, and K-NN, FCM and K-NN performs better in case of light abnormalities while ANFIS in case of dark abnormalities. Hassan Khojlanlou [9] proposed a technique which integrates region-based and contour based methods for improving precision of brain tumor segmentation. He uses fuzzy clustering technique, morphological operation and deformable model to implement proposed work. Logeswari [10] proposed brain tumor detection and segmentation method in which he divides proposed work in two phases—one is pre-processing and in second phase HSOM (Hierarchical self-organizing map) is applied. In last a brief comparison with other classifiers is done. Shasidhar [11] proposed a modified FCM algorithm for segmenting brain tumor. MFCM improves segmentation efficiency and computational time required for convergence. Rami J. Oweis [12] proposed pixel classification of medical image using neuro fuzzy approach, which rely on spatial properties of the image and uses multiscaled representation of image. S. Murugavalli [13] proposed a neuro fuzzy approach to detect tissues like white matter, gray matter, csf and tumor. He used HSOM and FCM algorithm to classify images. He gave a comparative study between SOM-K-mean, SOM-Fuzzy, HSOM-K-mean and HSOM-fuzzy. M. Sharma [14] proposed a technique for brain tumor augmentation using ANFIS and Genetic algorithm. She improved existing methods in terms of accuracy, sensitivity, specificity by using proposed algorithm. From previous work study it is analyzed that there are various combination of algorithms used for segmenting brain tumor. Proposed research work uses ANFIS, Genetic, FCM algorithms from them which strengthen individual weakness. The organization of the rest of this paper is as follows: Section 2 presents proposed work methodology, Section 3 discusses results. Finally, we present our conclusion and future scope in Section 4.

2. PROPOSED WORK

The proposed work is divided in to five steps: Adaptive Histogram Equalization, Morphological Erosion, Edge Detection (Canny edge detection), Feature Extraction using GLCM, feature selection using genetic algorithm and for Extracting tumor area (Fuzzy c-means segmentation) is used as shown in fig 2.1.
2.1 Raw MRI Image Database

Proposed research work uses Astrocytoma type of brain tumor images. Astrocytoma has four GRADES - Pilocytic Brain Tumor (GRADE-1), Low Grade Brain Tumor (GRADE-2), Anaplastic Brain Tumor (GRADE-3), Gieloblastoma Brain Tumor (GRADE 4). Pilocytic and Low Grade brain tumor came under the category of Benign Brain Tumor while Anaplastic and Gieloblastoma Brain Tumor are Malignant type of Brain Tumor.

2.2 Adaptive Histogram Equalization

First step of proposed work is to preprocess raw MRI images. Preprocessing is the basic step to remove noise. Performance of further steps relies on preprocessing. Proposed work uses CLAHE (Contrast limited adaptive histogram equalization) for improving contrast of MRI images. CLAHE is a contrast enhancement algorithm proposed by Karel Zuiderveld[14]. CLAHE does not work on entire image. Instead it works on small portion of image. Finally output of small areas are combined using bilinear interpolation method. CLAHE is applied using function adapthisteq which is built-in function in MATLAB 7.0 where I represent input image and A is resulting image.

MATLAB: A = adapthisteq(I)

2.3 Morphological Erosion

This is used as a image preprocessing tools to sharpen regions and to fill gaps of binarized image. There are four basic morphological operations are defined like dilation, erosion, opening and closing[15]. Here, proposed research work uses morphological erosion. In erosion every pixel which touches background pixel is converted in to background pixel. Objects become smaller after applying erosion. Mathematically erosion can be represented as,

\[(A \ominus B)(x) = \{ x \in X, x = a + b : a \in A b \in B \} \quad (1)\]
Where \( A \) is matrix of binary image and \( B \) is mask

\[
\text{MATLAB: } BW2 = 
\]

Where \( SE \) is structuring element

### 2.4 Edge Detection

Edge detection algorithms are used to find sudden changes in an image and also reduces amount of data by filtering useless information. There are various edge detection algorithm like sobel, prewit, LoG, Canny. This proposed research work uses Canny edge detection algorithm [16] because it performs better under all conditions [17][18].

\[
\text{MATLAB: EDGE (IMAGE, ’Canny’)};
\]

### 2.5 Feature Extraction

For an image, features are the characteristics of the objects. Texture is the main feature of an image. Feature extraction is the process of extracting some features from the pre-processed image. These features are used for segmenting image. Image segmentation can be done in two ways: stastical approach and structured approach. Proposed research work uses stastical approach. There are several stastical techniques for measuring texture such as co-occurrence matrix, Fractals, Gabor filters, wavelet transform [19]. Proposed research work uses Gray Level Co- occurrence Matrix (GLCM) .GLCM captures numerical feature values using spatial relationship among neighborhood pixels features. These numerical feature values are used for further comparing and classifying features [19].GLCM extract 20 texture features (Autocorrelation, Contrast, Correlation, Cluster Prominence, Cluster Shade, Dissimilarity, Energy, Entropy, Homogeneity, Maximum probability , Sum of squares, Variance, Sum average,, Sum variance , Sum entropy, Difference variance , Difference entropy ,information measure of Correlation, Information measure of correlation2 Inverse difference (INV), Inverse difference normalized (INN) Inverse difference moment normalized )which are calculated using function available in MATLAB 7.0 for a given image:

\[
\text{GLCM2 = graycomatrix(image,’Offset’,[2 0;0 2])}
\]

Where, image is used for grey scale image. Table 2.5.1 shows feature extracted for an image using function available in MATLAB

<table>
<thead>
<tr>
<th>Feature No</th>
<th>Feature Name</th>
<th>Feature Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>autocd</td>
<td>43.1530</td>
</tr>
<tr>
<td>2</td>
<td>contrd</td>
<td>1.8692</td>
</tr>
<tr>
<td>3</td>
<td>corrpd</td>
<td>0.1392</td>
</tr>
<tr>
<td>4</td>
<td>cpronmd</td>
<td>34.6933</td>
</tr>
<tr>
<td>5</td>
<td>cshad1</td>
<td>5.2662</td>
</tr>
<tr>
<td>6</td>
<td>energd</td>
<td>0.1233</td>
</tr>
<tr>
<td>7</td>
<td>Dissid</td>
<td>0.6877</td>
</tr>
<tr>
<td>8</td>
<td>entrod</td>
<td>2.6980</td>
</tr>
<tr>
<td>9</td>
<td>homopd</td>
<td>0.65645</td>
</tr>
<tr>
<td>10</td>
<td>maxprd</td>
<td>0.6411</td>
</tr>
<tr>
<td>11</td>
<td>sosvhd</td>
<td>0.1973</td>
</tr>
<tr>
<td>12</td>
<td>savghd</td>
<td>44.9329</td>
</tr>
<tr>
<td>13</td>
<td>svarhd</td>
<td>13.2626</td>
</tr>
<tr>
<td>14</td>
<td>senthd</td>
<td>133.5676</td>
</tr>
<tr>
<td>15</td>
<td>dvarhd</td>
<td>1.8188</td>
</tr>
<tr>
<td>16</td>
<td>denthd</td>
<td>1.8927</td>
</tr>
<tr>
<td>17</td>
<td>inf1hd</td>
<td>1.2145</td>
</tr>
<tr>
<td>18</td>
<td>inf2hd</td>
<td>-0.0322</td>
</tr>
<tr>
<td>19</td>
<td>indncd</td>
<td>0.2863</td>
</tr>
<tr>
<td>20</td>
<td>idmnncd</td>
<td>0.9107</td>
</tr>
</tbody>
</table>

### 2.6 Feature Selection using Genetic Algorithm and ANFIS (adaptive network fuzzy inference System)

Feature selection helps to reduce the features which improve the prediction accuracy and also computation time is reduced. The main aim of feature selection is to select those features which are relevant and informative. Mostly search procedures are used for feature sub set selection. Poplarly used feature selection algorithms are Sequential forward Selection, Sequential
Backward selection, Genetic Algorithm and Particle Swarm Optimization. Proposed research work uses Genetic algorithm for selecting features obtained from GLCM (Grey Level Co-Occurrence Matrix). Genetic algorithm is a popular soft computing technique which is used for feature subset selection. There are several approaches which follow Genetic algorithm to solve feature subset selection problem [20]. Genetic Algorithm is a very robust technique but computational time taken by it is high when no. of features are large. So, a filter approach is used with Genetic Algorithm. Proposed research work uses ANFIS (Adaptive neural network fuzzy inference System) as filtering tool for Genetic Algorithm. Genetic Algorithm based feature subset selection algorithm which uses ANFIS for fitness function follows following step:

1. GLCM extracts 20 texture features. Mark them as Feature 1, Feature 2 and so on.
2. Following Parameters For Genetic Algorithm are set:
   - Population Size: 10
   - Maximum Chromosome Length: 7
3. Initial Subset of 7 features is selected randomly from all possible solution sub space. These 7 features are represented by binary string (11010110101000000000), where 1 shows presence of feature and 0 absence of feature i.e. Feature no1, 2, 4, 5,6,7,9 are selected initially

\[ \text{Fitness (f)} = \sum_{i=1}^{20} \left( \frac{TP}{TP+FN} + \frac{TN}{TN+FP} + \frac{TP+TN}{TP+TN+FP+FN} \right) \]  

Where Fitness (f) means fitness value for a subset, True Positive (TP) means Both training algorithm and testing algorithm results are positive. True Negative (TN) is Both training algorithm and testing algorithm results are negative, False Positive (FP) is training algorithm result is positive and testing algorithm are negative and False Negative (FN) is Training algorithm result is negative and Testing algorithm results are positive.
4. Fitness Function: Fitness function decides the success of Genetic Algorithm. For proposed research work Fitness function is determined by ANFIS

\[ \text{Best Feature 

 Subset = Fitness(f) - Fitness_{max} } \]  

If the difference between two is 0 to 15 then that particular feature subset is considered as best feature subset. Otherwise mutation and crossover is done to generate new population. Go to step 4 again

2.5 Fuzzy C-Means Segmentation

Clustering is a mathematical tool that groups the data with similar feature vector in to one group and dissimilar data in to different group. In hard clustering, a feature vector can be a part of only one cluster. Whereas, in Fuzzy clustering feature vector can belong to more than one cluster but have different membership degrees (between 0 and 1).

Steps of Fuzzy C-means Algorithm:

Let \( X = \{x_1, x_2, x_3, ..., x_n\} \) represents feature vector and \( V = \{v_1, v_2, v_3, ..., v_c\} \) represents set of centers of cluster.
1) Randomly select ‘c’ cluster centers.
2) Fuzzy membership ‘µij’ is calculated using:

\[ \mu_{ij} = \frac{1}{\sum_{k=1}^{c}\left(d_{ij}^{m}\right)^{1/m}} \]  

(3)

3) Then, fuzzy centers 'v_j' are calculated using:

\[ V_j = \frac{\sum_{i=1}^{n}(\mu_{ij}^m x_i)}{\sum_{i=1}^{n} \mu_{ij}^m} \]  

(4)

4) Step 2) and 3) are repeated either minimum 'J' value of 'J' is achieved or ||U^{(k+1)} - U^{(k)}|| < \beta . Where ‘k’ represents iteration step and ‘\beta’ is the termination criterion having value between [0, 1].

"U = (\mu_{ij})_{n \times c}" represents fuzzy membership matrix. ‘J’ represents objective function.

Four clusters are formed against each Grade of Astrocytoma (I to IV). Proposed research work segment MRI brain image with all of its four GRADES. When a new test image is tested to know type of tumor. Firstly, image is pre-processed, then feature are extracted and compare with the trained images. Best match is the solution of proposed system.

3. Results

![Fig.3.1 Main GUI](image1)

![Fig.3.2 Process of image segmentation](image2)

Fig.3.1 shows the main graphical interface of proposed work. To test a image press load input then browse image which is to be test and then press process to segment image. Fig.3.2 shows process followed by proposed work. Tumor part is highlighted by red. Fig.3.3; Fig.3.4 shows the result for different image.
4. Conclusion and Future work

Performance of different image segmentation algorithm can be compared with following parameters:

True Positive (TP): Both Proposed Segmentation algorithm and manual segmentation by doctor’s results are positive

True Negative (TN): Both Proposed Segmentation algorithm and manual segmentation by doctors results are negative

False Positive (FP): Proposed Segmentation algorithm result is positive and manual segmentation results by doctors are negative.

False Negative (FN): Proposed Segmentation algorithm result is negative and manual segmentation by doctor’s results is positive.

Sensitivity = TP/ (TP+FN) *100%

Specificity = TN/ (TN+FP) *100%

Accuracy = (TP+TN)/ (TP+TN+FP+FN)*100 %

Table 4.1: Comparison of classification performance for the proposed technique and recently other work

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCM</td>
<td>96%</td>
<td>93.3%</td>
<td>86.6%</td>
</tr>
<tr>
<td>ANFIS+ Genetic</td>
<td>95%</td>
<td>93%</td>
<td>90%</td>
</tr>
<tr>
<td>K-Mean+ FCM</td>
<td>80%</td>
<td>93.12%</td>
<td>83.3%</td>
</tr>
<tr>
<td>Proposed (ANFIS+ Genetic+ FCM)</td>
<td>96.6%</td>
<td>95.3%</td>
<td>98.67%</td>
</tr>
</tbody>
</table>

Fig.4.1 Comparison graph of classification performance for the proposed technique and recently other work. A comparative study is shown in Table 4.1 and Fig.4.1. There are three performance measure to compare these algorithm i.e. Sensitivity, Accuracy, Specificity. Proposed algorithm shows improved accuracy, sensitivity, specificity from FCM, ANFIS plus Genetic, K-mean plus FCM. This proposed algorithm can be extended to segment other tumor types (Glioma, Metastases) also.

References


