Brain Tumor Segmentation by FCM and Enhancement by ANN, using GLCM Based Feature Extraction

Salabha Varghese

Sree Narayana Gurukulam College of Engineering, Kerala, India

Abstract: Brain Tumor is an uncontrolled growth of tissues in brain. Robust segmentation of brain tissues is very important for the detection and diagnosis of brain tumor. Clustering techniques could be best utilized by combining them with other techniques. In this paper the idea is to combine two metaphors: Neural Network and Fuzzy Logic. The Neuro-Fuzzy system combines the learning power of Artificial Neural Network system and explicit knowledge representation of fuzzy inference system. The proposed system consists of four stages: Pre-processing of various brain images, Feature extraction using Gray Level Co-occurrence Matrix (GLCM) and classification and segmentation of brain images through Hybrid Neuro-Fuzzy System.

Keywords: Clustering, Neural Network, FCM, Neuro-Fuzzy, GLCM

1. Introduction

Brain tumor is an abnormal growth of cells inside the skull. Tumors can damage the normal brain cells by producing inflammation, exerting pressure on parts of brain and increasing pressure within the skull. Accurate image segmentation is of utmost importance in diagnosis such deadly, because if not properly detected and diagnosed, these diseases may result in the death of the person.

In the past few decades, we can see that the death due to brain tumor is increasing at a faster rate. This is mainly due to the improper detection and diagnosis. Recent surveys too add to this fact.

Several methods are used to detect brain tumor. These methods mainly relay on image processing techniques. Mainly used image segmentation techniques are Edge based, region based and pixel based techniques. Edge based techniques like thresholding has certain disadvantages such as it ignores the cells whose threshold value is approaching the chosen threshold. The process of selecting the appropriate threshold is a tedious task done manually and the threshold varies from image to image. In region growing techniques, the region is grown based on homogeneity between the cells of one region with another. This method depends on the selection a seed pixel and may therefore lead to wrong results if the initial seed point is not chosen correctly. Pixel based clustering is another important segmentation technique. K-means clustering could be used for efficiently clustering tumor cells. But, this method is sensitive to outliers and depends on the number of clusters chosen at the beginning. Therefore, clustering techniques could be best utilized by combining them with other techniques.

This paper presents a hybrid approach to detect the shape and size of the tumor region. It also uses the power of artificial neurons to create a model that mimics the real neurons, and thus helps in the diagnosis of the disease.

2. Related Works

The image segmentation is entailed with the division or separation of the image into regions of similar features. Time consumption during the segmentation of brain tumor from magnetic resonance imaging is a crucial drawback. The segmentation & edge detection approaches fall mainly under 3 categories. These are as follows: 1) Thresholding approaches, 2) Region growing approaches, 3) Histogram based approaches, 4) Clustering approaches. Several authors suggested various algorithms for segmentation.

Jaskirat Kaur, Sunil Agrawal & Renu Vig.’s [5] paper presented thresholding and edge detection being one of the important aspects of image segmentation comes prior to feature extraction and image recognition system for analyzing images. It helps in extracting the basic shape of an image, overlooking the minute unnecessary details. In this paper using image segmentation (thresholding and edge detection) techniques different geo-satellite images, medical images and architectural images are analyzed. To quantify the consistency of our results error measure is used.


J.Vijay and J.Subhashini [8] proposed a method by keeping the clustering methods to segment tumor region. The unsupervised, K means algorithm was considered as the backbone of the method. Clustering starts by determining k initial central points. It then groups each image pixel under the central point it is closest to. The limitations of K-means clustering are many iterative rounds may be required and may prove good only in detecting spherical clusters.
J. Selvakumar, A. Lakshmi and T. Arivoli [7] proposed a method to overcome the above drawbacks. This paper provides a computer aided method for segmentation (detection) of brain tumor based on the combination of two algorithms. This method allows the segmentation of tumor tissue with accuracy and reproducibility comparable to manual segmentation. In addition, it also reduces the time for analysis. At the end of the process the tumor is extracted from the MR image and its exact position and the shape also determined.

A combined elastic atlas registration with statistical classification has been introduced by Warfield [15]. A modified spatial atlas for classification including prior probabilities for brain tumor and edema is presented by Marcel Prastawa in his research work [9]. For biomedical image classification problems the unsupervised learning based Self Organizing Feature Map (SOFM) algorithms guarantee excellent classification accuracy in the classification of images having brain tumor [6]. For unsupervised pattern classification of medical images, especially in the detection of abnormal tissues, Hopfield neural networks (HNN) have been proven as a robust and efficient method.

The ultimate goal implementing a hybrid system is to overcome the problems related to imprecise information present in a MRI image such as greyness of a pixel and ambiguity, uncertain interpretation of a scene and geometrical segmentation of the image. The neuro fuzzy approach exploits the learning capability and descriptive power of the systems, Hence it provides good accuracy and high interpretability. Rami J. Oweis and Muna J. Sunna [11] has been proposed segmentation of images using neuro fuzzy model. Neuro fuzzy tools based image segmentation has also been implemented by Mausumi Acharyya [2].

3. Proposed Method

The proposed system introduces a mechanism by which brain tumour could be detected with pin-point perfection. The system overcome the drawbacks of the existing system. It has mainly four modules:

- Pre-processing
- Feature extraction
- Neural network classification
- Segmentation
- Tumor size calculation

Pre-processing provides the input image (in the required format) for the next stage. The next stage is feature extraction. Here, the features of the trained images like intensity; co-variance, entropy etc. are extracted and stored in the database.
distance \( d \), and angle \( \theta \). The sum of the number of times that the pixel with value \( i \) occurred in specified spatial relationship to a pixel with value \( j \) in the input image, is represented by each element \((i,j)\) in the resultant GLCM. In a GLCM matrix, there are same number of gray levels \( G \) in image, as the numbers of rows and columns in the matrix. The relative frequency is represented by the matrix element \( P(i, j \mid \Delta x, \Delta y) \) separated by a pixel distance \((\Delta x, \Delta y)\). Matrix element may also be expressed by \( P(i, j \mid d, \theta) \) which contains the 2nd order probability values that is associated with the changes between gray levels \( i \) and \( j \) at a distance of \( d \) and an angle of \( \theta \).

Following equations are used to compute the different textural features that can be used to train the Neuro-Fuzzy classifier.

### Homogeneity (Angular Second Moment)
\[
\text{ASM} = \sum_{i=1}^{G} \sum_{j=1}^{G} P(i,j)^2
\]

### Inverse Difference Moment
\[
\text{IDM} = \sum_{i=1}^{G} \sum_{j=1}^{G} \frac{1}{1+|i-j|^2} P(i,j)
\]

### Contrast
\[
\text{Contrast} = \sum_{i=1}^{G} \sum_{j=1}^{G} (i-j)^2 P(i,j), \quad |i-j| = n
\]

### Variance
\[
\text{Variance} = \sum_{i=1}^{G} \sum_{j=1}^{G} (i-\mu)^2 P(i,j)
\]

### Correlation
\[
\text{Correlation} = \frac{\sum_{i=1}^{G} \sum_{j=1}^{G} (i-\mu)(j-\mu) P(i,j)}{\sum_{i=1}^{G} \sum_{j=1}^{G} (i-\mu)^2 P(i,j) \cdot \sum_{i=1}^{G} \sum_{j=1}^{G} (j-\mu)^2 P(i,j)}
\]

### Entropy
\[
\text{Entropy} = -\sum_{i=1}^{G} \sum_{j=1}^{G} P(i,j) \log (P(i,j))
\]

### Inertia
\[
\text{Inertia} = \sum_{i=1}^{G} \sum_{j=1}^{G} (i-j)^2 P(i,j)
\]

### Cluster Shade
\[
\text{Shade} = \sum_{i=1}^{G} \sum_{j=1}^{G} (i + j - \mu_x - \mu_y)^2 P(i,j)
\]

### Cluster Prominence
\[
\text{Prom} = \sum_{i=1}^{G} \sum_{j=1}^{G} (i + j - \mu_x - \mu_y)^2 P(i,j)
\]

### Sum Average
\[
\text{Aver} = \sum_{i=1}^{G} \sum_{j=1}^{G} P(i,j)
\]

### Sum Entropy
\[
\text{Sent} = -\sum_{i=1}^{G} \sum_{j=1}^{G} P(i,j) \log (P(i,j))
\]

### Difference Entropy
\[
\text{Dent} = -\sum_{i=1}^{G} \sum_{j=1}^{G} P(i,j) \log (P(i,j))
\]

The values of these GLCM features decide whether the input brain image has any abnormality or not. For Brain MRI images; normal images must have contrast value in the range of 0.7000 to 0.7550 for GLCM features. Input images which are outside this range are classified as abnormal images by the classifier.

### 3.3 Artificial Neural Network Classifier

In the proposed methodology, we employ artificial neural networks to classify the brain cells as normal and abnormal cells. The neural network models try to emulate the architecture and information representation schemes of the human brain. Artificial Neural Networks are parallel adaptive networks of simple non-linear computing elements called neurons which are intended to abstract and model some of the functionality of human nervous system in an attempt to partially capture some of its computational strengths. It can be any computable function, by the appropriate selection of the network topology and weights values. Neural networks offer a number of advantages, including requiring less formal statistical training, ability to implicitly detect complex nonlinear relationships between dependent and independent variables, ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms.

3.4 Segmentation

The purpose of image segmentation is to distinguish the meaningful objects in an image. It partitions an image into a set of homogeneous regions. Here, segmentation process deals with the extraction of tumour regions from ordinary brain cells. It is done by:

![Figure 3.3: (a) A typical artificial neural network](image1)

![Figure 3.3: (b) Neuro-fuzzy classifier](image2)
• K-means clustering algorithm and
• Fuzzy C-means algorithm

![Schematic diagram of segmentation process](image-url)

**3.4.1 K-Means Clustering Technique**

The K-means algorithm is well known for its efficiency in clustering large data sets. It is an unsupervised learning algorithm. Clustering the image is the process of grouping the pixels according to some characteristics, like intensity, texture etc. In the K-means algorithm, initially we have to define the number of clusters k or it could be set to a predefined number.

K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean.

Given a set of observations \((x_1, x_2, \ldots, x_n)\), where each observation is a \(d\)-dimensional real vector, the K-means clustering aims to partition the \(n\) observations into \(k\) sets \(\{S_1, S_2, \ldots, S_k\}\)

\[
\arg\min_{\mu} \sum_{i=1}^{k} \sum_{x_j \in S_i} \|x_j - \mu_i\|^2
\]

where \(\mu_i\) is the mean of points in \(S_i\).

Given an initial set of \(k\) means \(m^{(1)}, \ldots, m^{(1)}\) the algorithm proceeds by alternating between two steps:

**Assignment step:** Assign each observation to the cluster whose mean yields the least within-cluster sum of squares (WCSS). Since the sum of squares is the squared Euclidean distance, this is intuitively the nearest mean.

**Update step:** Calculate the new means to be the centroids of the observations in the new clusters.

\[
m^{(t+1)}_i = \frac{1}{|S^{(t)}_i|} \sum_{x_j \in S^{(t)}_i} x_j
\]

The generalized algorithm initiates \(k\) cluster centroids by randomly selecting \(k\) feature vectors from \(X\). Later, the feature vectors are grouped into \(k\) clusters using a selected distance measure such as Euclidean. The clustering procedure stops only when all cluster centroids tend to converge. Similarity is measured by distance and defined by an \(N\) dimensional feature space. Feature distance calculation is based on features such as color or intensity and texture while spatial distance calculation is based on \(x\), \(y\) (width, height) coordinates.

![Flowchart of the algorithm](image-url)

Even though K-means is an efficient clustering algorithm, it has certain disadvantages like becoming sensitive to outliers, finding only spherical clusters and it often becomes a bottleneck when clusters have different sizes. Also, working only on numeric values prohibits it from being used to cluster real world data containing categorical values. To overcome these defects, the clustered image is passed through Fuzzy C-means algorithm.

**3.4.2 Fuzzy C-Means Algorithm**

The fuzzy logic is a way to processing the data by giving the partial membership value to each pixel in the image. The Fuzzy C-Means (FCM) algorithm uses fuzzy logic where each data point is specified by a membership grade between 0 and 1. The membership function defines the fuzziness of an image and also to define the information contained in the image.

It provides a fuzzy partition of the image by giving each pixel a degree of belonging to a given region. In the process of segmentation, cells which share common intensity values are grouped together. A representative membership value is chosen from each group. Moreover, it gives the natural representation of the behaviour of genes. Fuzzy logic is thus becoming a milestone in medical diagnosis.

**Algorithm**

1. Initialize \(M=M^{(0)}\) matrix, \(M^{(0)}\)
2. At k-step: calculate the centres vectors \(R(k)=[R_i]\) with \(M^{(k)}\)
\[ R_j = \frac{1}{N} \sum_{i=1}^{N} x_i \cdot M_{ij}^m \]

where \( m \) is the fuzzy parameter and \( N \) is the number of data points.

3. Update \( M^{(k)} \), \( M^{(k+1)} \)

\[ M_{ij}^{(k+1)} = \left( \frac{1}{C} \sum_{c=1}^{C} \left( \frac{x_i - c_j}{\delta} \right) \right)^{-\frac{1}{m}} \]

4. If \( \| M^{(k+1)} - M^{(k)} \| < \delta \), then STOP; otherwise return to step 2.

\( k \): no: iteration steps
\( \delta \): termination value = constant between 0 and 1
\( \delta \): membership cut off value

Finally, morphological operations like erosion and dilation are applied over the clustered image to obtain the segmented region accurately.

3.5 Tumor Size Calculation

For the accurate diagnosis of the disease, the exact tumor size should be known. In this step, the tumor area is calculated using the binarization method. The image is considered to have only two values either black or white (0 or 1). Here, 256x256 jpeg image is a maximum image size.

\[ image, I = \sum_{h=0}^{255} \sum_{w=0}^{255} [f(0) + f(1)] \]

Pixels = Width (W) x Height (H) = 256 x 256
\( f(0) \) = black pixel digit (0)
\( f(1) \) = white pixel digit (1)

No: of white pixels, \( P = \sum_{h=0}^{255} \sum_{w=0}^{255} f(1) \)

1 Pixel = 0.264 mm

Size of tumor, \( S = \left( \sqrt{P} \right) \times 0.264 \) mm²

P = no: of white pixels
W = width
H = height

Using the above equations, the size of the tumor region could be obtained and it could play an important role in future diagnosis of the patient.

4. Results

In the initial preprocessing step we use median filter to remove the noise in the input image. The filtered image is then given for neural network classification and the segmentation process relays on its result. In the segmentation process clustering techniques are used. And finally, area of the segmented tumor region is calculated. The results of the proposed method could be best illustrated by the screenshots below:

5. Conclusion

In this paper, an efficient method for the brain tumor detection is presented. It provides fast and robust results compared to the existing system. The input image chosen should be (preferably) taken from an MRI scan. The neuro-fuzzy approach overcomes the drawbacks of the neural network and fuzzy logic, which occurs when used alone. After the pre-processing stage, the features of the image are extracted using the GLCM technique. The images in the database are trained by the implemented artificial neural network classifier. The fuzzy logic models attempt to mimic human reasoning and the capability of handling uncertainty. Fuzzy logic gives the natural representation of the behaviour of genes, since fuzzy logic does not stick on crisp boundaries.

The proposed method is computationally feasible in case of time and large data sets. Its approach is simple and easy to understand since the pixel-based segmentation methods are simple and the computational complexity is relatively low and accuracy is high compared with other region-based or edge-based methods, the application is more practicable.

6. Future Scope

In future, the system could be redone with 4-D technologies and upcoming segmentation methods.
References


Author Profile

Salabha Varghese received B.Tech Degree in Computer Science and Engineering from Cochin University of Science and Technology, Kerala, India, in 2011. Currently pursuing M.Tech Degree in Computer Science and Engineering, under Mahathma Gandhi University, Kerala, India