Brain Tumor Segmentation Based on SFCM using Back Propagation Neural Network

Sake Anjali Durga[1], Ponakala Lakshmi[2], Pallabotula Bhagya Sri[3], Ravulapalli Anusha[4],
V. Purna Chandra Reddy[5]

[1,2,3,4] *B.Tech Students, Dept. Of ECE,VVIT, Nambur, Guntur District, A.P, INDIA.*
[5] Assoc. Prof., Dept. Of ECE, VVIT, Nambur, AP, INDIA.

---***---

Abstract – *Automatic defects detection in MR images is very important in many diagnostic and therapeutic applications. Because of high quantity data in MR images and blurred boundaries, tumor segmentation and classification is very hard. This work has introduced one automatic brain tumor detection method to increase the accuracy and yield and decrease the diagnosis time. The goal is classifying the tissues to three classes of normal, begin and malignant. In MR images, the amount of data is too much for manual interpretation and analysis. During past few years, brain tumor segmentation in magnetic resonance imaging (MRI) has become an emergent research area in the field of medical imaging system. Accurate detection of sizing and location of brain tumor plays a vital role in the diagnosis of tumor. The diagnosis method consists of four stages, pre-processing of MR images, future extraction, and classification. After histogram equalization of image, the features are extracted based on Dual Tree Complex wavelet transformation (DWT). In the last stage, Neural Network (NN) are employed to classify the normal and abnormal brain. An efficient algorithm is proposed for tumor detection based on the spatial Fuzzy C-Means Clustering.*

Key Words: Dual Tree Complex Wavelet Transformation (DTCWT), Magnetic Resonance Imaging (MRI), Spatial Fuzzy C Means (SFCM), Neural Network (NN), Very Deep Convolutional Networks.

1. INTRODUCTION

The unique identification of objects in an image would probably start with image processing techniques such as noise removal, followed by (low-level) feature extraction to locate lines, regions and possibly areas with certain textures.

An image is usually interpreted as a two-dimensional array of brightness values, and is most familiarly represented by such patterns as those of a photographic print, slide, television screen, or movie screen. An image can be processed optically or digitally with computer.

To digitally process an image, it is first necessary to reduce the image to a series of numbers that can be manipulated by the computer. Each number representing the brightness value of the image at a particular location is called a picture element, or pixel. A typical digitized image may have 512 x 512 or roughly 250,000 pixels, although much larger image s are becoming common. Once the image has been digitized, there are three basic operations that can be performed.

These operations taken singly or in combination, are the means by which the image is enhanced, restored, or compressed. An image is enhanced when it is modified so that the information it contains is more clearly evident, but enhancement can also include making the image more visually appealing.
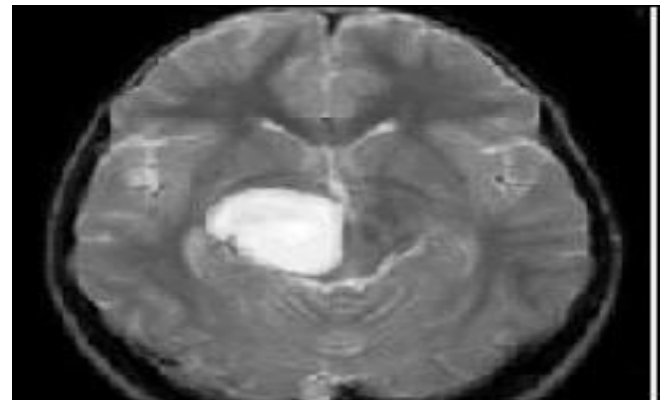
2. CLASSIFICATION OF IMAGES & PROCESSING:

Figure 1: Input MRI image.

Figure 2: Discrete Wavelet Transform Image.

There are 3 types of images used in Digital Image Processing. They are:
1. Binary Image.
2. Gray Scale Image.
3. Color Image.

PREPROCESSING:

Image restoration is the operation of taking a corrupted/noisy image and estimating the clean original image. Corruption may come in many forms such as motion blur, noise, and camera misfocus. Image restoration is different from image enhancement in that the later is designed to emphasize features of the image that make the

image more pleasing to the observer, but not necessarily to produce realistic data from a scientific point of view. Image enhancement techniques (like contrast stretching or de-blurring by a nearest neighbor procedure.) provided by "Imaging Packages" use no a priori model of the process that created the image. With image enhancement noise can be effectively be removed by sacrificing some resolution, but this is not acceptable in many applications. In a fluorescence microscope resolution in the Z-direction is bad as it is. More advanced image processing techniques must be applied to recover the object. De convolution is an example of image restoration method. It is capable of increasing resolution, especially in the axial direction removing noise increasing contrast.

In the neighborhood of an edge, the real DWT produces both large and small wavelet coefficients. In contrast, the analytic CWT produces coefficients whose magnitudes are more directly related to their proximity to the edge. Here, the test signal is a step edge at n=no, x(n)=u(n-no).

3. CLUSTERING MODEL

Clustering can be considered the most important unsupervised learning problem, so it deals with finding a structure in a collection of unlabeled data. A cluster is therefore a collection of objects which are "similar" between them and are "dissimilar" to the objects belonging to other clusters.

4. FUZZY CLUSTERING MODEL

Fuzzy clustering plays an important role in solving problems in the areas of pattern recognition and Fuzzy model identification. A variety of Fuzzy clustering methods have been proposed and most of them are based upon distance criteria. One widely used algorithm is the fuzzy c-means (FCM) algorithm. It uses reciprocal distance to compute fuzzy weights. A more efficient algorithm is the new FCFM. It computes the cluster centers using Gaussian weights, uses large initial prototypes, and adds processes of eliminating, clustering and merging.in the following sections we discuss and compare the FCM algorithm and FCFM algorithm.

Spatial Fuzzy C-Means method incorporates spatial information, and the membership weighting of each cluster is altered after the cluster distribution in the neighborhood is considered. The first pass is the same as that in standard FCM to calculate the membership function in the spectral domain. In the second pass, the membership information of each pixel is mapped to the spatial domain and the spatial function is computed from that. The FCM iteration proceeds with the new membership that is incorporated with the spatial function. The iteration is stopped when the maximum difference between cluster centers or membership functions at two successive iterations is less than a least threshold value.

The Fuzzy C-Means algorithm was introduced by J.C.Bezdek. The idea of FCM is using the weights that minimize the total weighted mean square error.

$$J(W_{qk}, Z^{(k)}) = \sum_{(k=1,k)}\sum_{(k=1,k)}(W_{qk})||x^{(q)}-z^{(k)}||^2$$

$$\sum_{(k=1,k)}(W_{qk})=1$$

$$W_{qk} = (1/(D_{qk})^2)^{1/(p-1)}/\sum_{(k=1,k)}(1/(D_{qk})^2)^{1/(p-1)}, p>1$$

The FCM allows each feature vector to belong to every cluster with a Fuzzy true value (between 0 and 1). The algorithm

assigns a feature vector to a cluster according to the maximum weight of the feature vector over all clusters.

5. ALGORITHM FLOW

Initialize the Fuzzy weights: In order to comparing the FCM with FCFM, our implementation allows the user to choose initializing the weights using feature vectors or randomly. The process of initializing the weights feature vectors assigns the first K_{init} (user-given) feature vectors to prototypes then computes the weights.
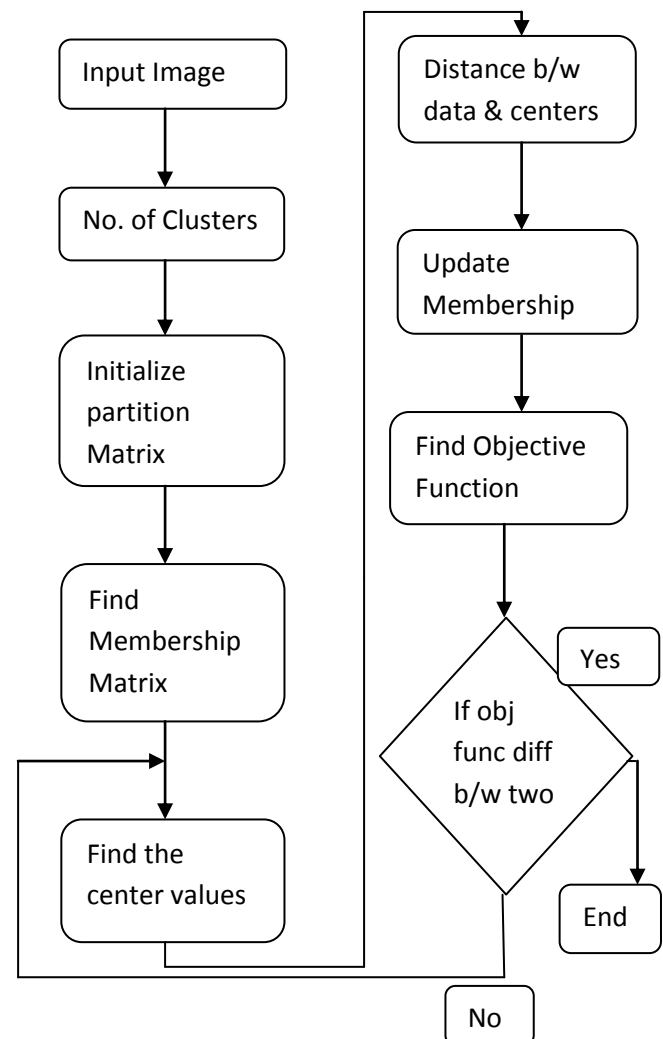

Figure 3: Algorithm flow for Fuzzy C-means

Standardize the weights over Q: During the FCM iteration, computed cluster centers get closer and closer. To avoid the rapid convergence and always grouping into one cluster, we use

$W[q,k]=(W[q,k]-W_{min})/(W_{max}-W_{min})$
Where, W_{max}=Maximum weight over the weights of all features.
        Wmin=Minimum weight over the weights of features.
Eliminating Empty Clusters : After the fuzzy clustering loop we add a step to eliminate the empty clusters. This step is put outside the fuzzy clustering loop and before calculation of modified XB validity. Without the elimination, the minimum distance of prototype pair used may be the

distance of empty cluster pair. We call the method of eliminating small clusters by passing zero to the process so it will only eliminate the empty clusters.

6. MORPHOLOGICAL PROCESS

Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. Morphological operations rely on the relative ordering of pixel values, not on their numerical values, and therefore are especially suited to the processing of binary images. Morphological operations can also be applied to grey scale images such that their light transfer functions are known and therefore their absolute pixel values are of no or minor interest.

Morphological techniques probe an image with a small shape or template called a **structuring element.** The structuring element is positioned at all possible locations in the image and it is compared with the corresponding neighborhood, of pixels. Some operations test whether the element "fits" within the neighborhood, while others test whether it "hits" or intersects the neighborhood:
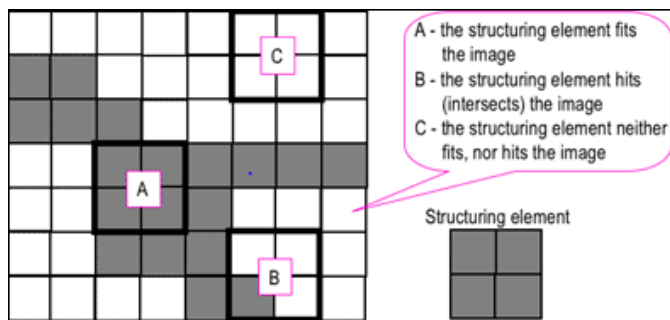

Figure 4: Probing of an image with a structuring element.

Erosion with small square structuring elements shrinks an image by stripping away a layer of pixels from both the inner and outer boundaries of regions. The holes and gaps between different regions become larger, and small details are eliminated:
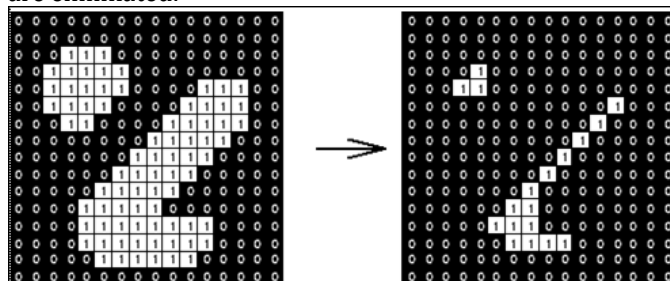

Figure 5: Erosion a 3x3 square structuring element.

7. FEATURE EXTRACTION AND NEURAL NETWORK
CO-OCCURRENCE MATRIX

Originally proposed by R.M. Haralick, the co-occurrence matrix representation of texture features explores the grey level spatial dependence of texture [2]. A mathematical definition of the co-occurrence matrix is as follows [4]:

- Given a position operator **P(i,j).**

- Let **A** be an n x n matrix.
- Whose element **A[i][j]** is the number of times that points with grey level(intensity),**g[i]** occur, in the position specified by **P,** relative to points with grey level **g[j].**
- Let **C** be the **n x n** matrix that is produced by

dividing A with the total number of point pairs that satisfy **P**. **C[i][j]** is a measure of the joint probability that a pair of points satisfying **P** will have values **g[i]**,**g[j]**.

- **C** is called a co-occurrence matrix defined by **P.**

```
1   2   1   3   4
2   3   1   2   4
3   3   2   1   1
```

Figure 6: Image example

	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	0	0	0	0	0
2	0	1	0	2	0	0	0	0
3	0	0	1	1	0	0	0	0
4	0	1	0	0	1	0	0	0
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Figure 7: Classical Co-occurrence matrix

Haralick proposed the following texture features:

1. Energy
2. Contrast
3. Correlation
4. Homogeneity

Energy: It is a gray-scale image texture measure of homogeneity changing, reflecting the distribution of image gray-scale uniformity of weight and texture.

$$E=\sum\sum p(x,y)^2$$

$P(x,y)$ is the GLCM

Contrast: Contrast is the main diagonal near the moment of inertia, which measure the value of the matrix is distributed and images of local changes in number, reflecting the image clarity and texture of shadow depth.

$$I=\sum\sum (x-y)^2 p(x,y)$$

Correlation Coefficient: Measures the joint probability occurrence of the specified pixel pairs.

Homogeneity: Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

$$Homogeneity=sum(sum(p(x,y)/(1+[x-y])))$$

8.NEURAL NETWORK:

Neural networks are predictive models loosely based on the action of biological neurons.

The selection of the name "neural network" was one of the great PR successes of the 20th century. It certainly sounds more exciting than a technical description such as "A network of weighted, additive values with nonlinear transfer functions". However, despite the name, neural networks are far from "thinking machines" or "artificial Lungs". A typical artificial neural network might have a hundred neurons. In comparison, the human nervous system is believed to have about 3×10^{10} neurons. We are still light years from "Data".

The original "Perception" model was developed by Frank Rosenblatt in 1958. Rosenblatt's model consisted of three layers, (1) a "retina" that distributed inputs to the second layer, (2)"association units" that combine the inputs with and trigger a threshold step function which feeds to the output layer, (3)the output layer which combines the values. Unfortunately, the use of a step function in the neutrons made the perceptions difficult or impossible to train. A critical analysis of perceptions published in 1969 by Marvin Minsky and Seymore Papert pointed out a number of critical weakness of perceptions. And for a period of time, interest in perceptions waned.

Interest in neural network was revived in 2986 when David Rumelharl, Geoffery Hinton and Ronald Williams published "Learning Internal Representations by Error Propagation". They proposed a multilayer neural network with nonlinear but differentiable transfer functions that avoided the pitfalls of the original perception's step functions. They also provided a reasonably effective training algorithm for neural network.

CONCLUSION

In summary, we propose a novel CNN-based method for segmentation of brain tumors in MRI images. We start by a pre-processing stage consisting of bias field correction, intensity and patch normalization. After that, during training, the number of training patches is artificially augmented by rotating the training patches, and using samples of HGG to augment the number of rare LGG classes. The CNN is built over convolutional layers with small 3x3 kernels to allow deeper architectures. In designing our method, we address the heterogeneity caused by multi-site multi-scanner acquisitions of MRI images using intensity normalization as proposed by Nyul´ et al. We show that this is important in achieving a good segmentation. Brain tumors are highly variable in their spatial localization and structural composition, so we have investigated the use of data augmentation to cope with such variability.

We studied augmenting our training data set by rotating the patches as well as by sampling from classes of HGG that were underrepresented in LGG. We found that data augmentation was also quite effective, although not thoroughly explored in Deep Learning methods for brain tumor segmentation. Also, we investigated the potential of deep architectures through small kernels by comparing our deep CNN with shallow architectures with larger filters. We found that shallow architectures presented a lower performance, even when using a larger number of feature maps. Finally, we verified that the activation function LRELU was more important than LRELU in effectively training our CNN. We evaluated the proposed method in BRATS 2013 and 2015 databases.

Concerning 2013 database, we were ranked in the first position by the online evaluation platform. Also, it was obtained simultaneously the first position in DSC metric in the complete, core, and enhancing regions in the Challenge data set. Comparing with the best generative model, we were able to reduce the computation time approximately by ten-fold. Concerning the 2015 database, we obtained the second position among twelve contenders in the on-site challenge. We argue, therefore, that the components that were studied have potential to be incorporated in CNN-based methods and that as a whole our method is a strong candidate for brain tumor segmentation using MRI images.

ACKNOWLEDGEMENT

REFERENCES

[1] *Ardizzone.E, Pirrone.R, and Orazio.O.G, "Fuzzy C-Means Segmentation on Brain MR Slices Corrupted by RF-Inhomogeneity," In Proc. The 7th international workshop on Fuzzy logic and Applications: Applications of Fuzzy Sets Theory, WILF '07, Springer-Verlag, pp:378-384,2007.*

[2] *Bianrgi.P.M, Ashtiyani.M, and Asadi.S, "MRI Segmentation using Fuzzy C-Means Clustering Algorithm basis Neural Network," In proc. ICTT A 3rd international conference on information and communication technologies: From theory to applications, pp:1-5, 2008.*

[3] *Sikka.K,Sinha.N, Singh.P.K, and "A fully automated algorithm under modified FCM framework for improved brain MR image segmentation," Magn. Reson, Imag, vol.27,pp. 994—1004, Jul. 2009.*

[4] *Xiao.K, Ho.S.H, and Bargiela.B, "Automatic Brain MRI Segmentation Scheme based on Feature Weighting Factors Selection on Fuzzy C-Means Clustering Algorithms with Gaussian smoothing, "International Journal of Computational Intelligence in Bioinformatics and systems Biology(3):316-331, 2009.*

[5] *B. Menze et al., "The multimodal brain tumor image segmentation benchmark (brats)," IEEE Transactions on Medical Imaging, vol. 34, no. 10, pp. 1993–2024, 2015.*