Performance Evaluation of Basic Segmented Algorithms for Brain Tumor Detection

[1]Suchita Yadav , [2]Sachin Meshram
[1]*(Department of ETC / Chouksey Engineering College, Bilaspur India)*
[2]*(Department of ETC / Chouksey Engineering College, Bilaspur India)*

Abstract*: In the field of computers segmentation of image plays a very important role. By this method the required portion of object is traced from the image. In medical image segmentation, clustering is very famous method . By clustering, an image is divided into a number of various groups or can also be called as clusters. There are various methods of clustering and thresholding which have been proposed in this paper such as otsu , region growing , K Means , fuzzy c means and Hierarchical self organizing mapping algorithm. Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method (developed by Dunn in 1973 and improved by Bezdek in 1981) is frequently used in pattern recognition. As process of fuzzy c mean is too slow, this drawback is then removed. In this paper by experimental analysis and performance parameters the segmentation of hierarchical self organizing mapping method is done in a better way as compared to other algorithms. The various parameters used for the evaluation of the performance are as follows: segmentation accuracy (Sa) , area (A), rand index (Ri),and global consistency error (Gce) .*

Keywords - *area (A), Fuzzy C means, global consistency error (Gce) , HSOM, K means , Otsu , rand index (Ri), Region Growing , segmentation accuracy (Sa) , and variation of information (Vi).*

I.   Introduction

The brain is a very important part of the body and is one of the special organ . The brain consists of large number of cells which grow on increasing from third month upto the seven years of age. Each cell has its own special function. When some of the cells in the body grow in an orderly way to generate new cells then the body is healthier. But if cells grow irregularly then various cells which are in excess form a cluster or mass of tissue known as tumor. Brain tumor is one of the most common and deadly diseases in the world. Detection of the brain tumor in its early stage is the key of its cure [1]. Brain tumors may be benign or malignant[2].

The term gray level is often used to refer to the intensity of the monochrome images. In the field of medical, segmentation has wide application . Color images are formed by a combination of individual 2-D images. An image may be continuous with respect to the x and y coordinates and also in amplitude. A continuous image is converted to digital image which requires that the coordinates as well as the amplitude be digitized. By using image segmentation, the image can be divided into various parts such as mutually exclusive and exhausted regions. Image segmentation is an important and challenging factor in the medical image segmentation [3]. The ultimate aim in a large number of image processing applications is to extract important features from the image data, from which a description, interpretation, or understanding of the scene can be provided by the machine .Image segmentation is too complex to be used in image processing and is very sensitive and important component when image is being analyzed. In this paper various methods of image segmentation are used, mentioned as follows : Thresholding such as Region growing, Otsu,Clustering such as k means, Fuzzy cmeans and Hierarchical Self Organizing mapping [4].It utilized the HSOM, thresholding and clustering methods to identify which type of brain tumor suffered by patient regarding to the image of brain tumor from the Magnetic Resonance Imaging (MRI) and scan as inputs for the network and other methods.

II.   Problem Description

The required area   of an image is exploited by using the undesired component, atmospheric interference. So rather than analyzing the original image, the image segmentation technique is used.Various experiments with published benchmarks are required for this research field to progress [4] .The drawback which occurs in this paper is that image is divided into number of segmentation [5] .This drawback is overcome by selecting an appropriate model for segmentation and then modified with reduced computational time and output is of high quality.

III.   Segmentation Of Image By Thresholding And Clustering Methods

Clustering can be taken as a process of partitioning or grouping a given portion which is an unlabeled pattern into a large number of clusters such that a group is assigned to similar patterns. Clustering is used for pattern recognition in image processing, and usually requires a high volume of computation. Thresholding is

the simplest method of image segmentation. From a grayscale image, thresholding can be used to create binary images. During the thresholding process, individual pixels in an image are marked as "object" pixels if their value is greater than some threshold value (assuming an object to be brighter than the background) and as "background" pixels otherwise. This convention is known as threshold above. Variants include threshold below, which is opposite of threshold above; threshold inside, where a pixel is labeled "object" if its value is between two thresholds; and threshold outside, which is the opposite of threshold inside. Typically, an object pixel is given a value of "1" while a background pixel is given a value of "0." Finally, a binary image is created by coloring each pixel white or black, depending on a pixel's labels. The major drawback to threshold-based approaches is that they often lack the sensitivity and specificity needed for accurate classification.

1 Fuzzy C Means Algorithm

The goal of a clustering analysis is to divide a given set of data or objects into a cluster, which represents subsets or a group. The partition should have two properties: 1. Homogeneity inside clusters: the data, which belongs to one cluster, should be as similar as possible. 2. Heterogeneity between the clusters: the data, which belongs to different clusters, should be as different as possible. Clustering is a process to obtain a partition P of a set E of N objects Xi (i=1, 2,…, N), using the resemblance or dissemblance measure, such as a distance measure d. A partition P is a set of disjoint subsets of E and the element Ps of P is called cluster and the centers of the clusters are called centroid or prototypes. Many techniques have been developed for clustering data. In this report c-means clustering is used. It's a simple unsupervised learning method which can be used for data grouping or classification when the number of the clusters is known[4]. It consists of the following steps:
Step 1: Choose the number of clusters - K
Step 2: Set initial centers of clusters c1, c2… ck
Step 3: Classify each vector x [x , x ,....x] T into the closest centre c_i by Euclidean distance measure $\|x_i - c_i\|$ = min $\| x_i - c_i\|$
Step 4: Recomputed the estimates for the cluster centers ci. Let ci = [ci1 ,ci2 ,....cin] T
cim be computed by, cim = Σxli ∈ Clter(Ixlim) Ni .Where, Ni is the number of vectors in the ith cluster. Step 5:If none of the cluster centers changes in step 4 stop; otherwise go to step 3.

2 K Means

It is also one of the clustering method and is very famous because it is simpler and easier in computation. It is the simplest unsupervised learning algorithms that solve the well known clustering problem. It classifies the input data points into multiple classes based on their intrinsic distance from each other. The algorithm assumes that the data features form a vector space and tries to find natural clustering in them[4]. The algorithm which follows for the k-means clustering is given below: The cluster centers are obtained by minimizing the objective function :

$x_i \in s_i$ .... (1)

*1.*Initialize the centroids with k random values. 2. Repeat the following steps until the cluster labels of the image do not change anymore. 3. For each data point, we calculate the Euclidean distance from the data point to the mean of each cluster :

$C(i) = arg\ min \| x(i) - \mu_j \|^2$ ……………………..........(2)

If the data point is not closest to its own cluster, it will have to be shifted into the closest cluster. If the data point is already closest to its own cluster, we will not shift it. 4. Compute the new centroid for each of the clusters. Where *k* is a parameter of the algorithm (the number of clusters to be found), i iterates over the all the intensities, *j* iterates over all the centroids and μ_i are the centroid intensities.

3 Region Growing

Region growing comes under region-based segmentation method [5]. In this an image is partition into regions. There is predefined criteria where there are k clusters Si , i= 1,2,….,k and μ_i is the centroid or mean point of all the points, according to which groups of pixels are formed or groups of sub-regions are formed into larger regions as a result region growing takes place If suppose there is no priori information, then the procedure is to compute at every pixel the same set of properties that ultimately will be used to compute at every pixel the same set of properties that ultimately will be used to assign pixels to regions during the growing process.

4 Otsu Method

In image processing, Otsu's method is used for thresholding [6] of image by histogram method which is done automatically. The gray level image is converted to a binary image. In this method, the image on which this thresholding process has to be done is done in such a way that there are two models of histogram, known as bimodal histogram of the image. In other words, there are two classes of pixels one is background and other is

foreground. In image segmentation, the sub-division of image is done by using intensity of background and object. So both the different regions are then distinguished by a suitable value of threshold value. It consists of two classes within class variance and between class variance. In Otsu's method the value the threshold is chosen such that it minimizes the intra-class variance, defined a weighted sum of variances of the two classes :

$$\sigma^2_{\omega}(t) = \omega_1(t)\,\sigma_1^2(t) + \omega_2(t)\,\sigma_2^2(t) \qquad \text{……………………………..(3)}$$

$\omega 1$ & $\omega 2$ are weights, probabilities of the two classes separated by a threshold and variances of these classes.In this, threshold operation is done in such a way that an image is sub-divided into two classes R0 and R1 at gray level t. Let partitioning of the pixels is done between objects and background. Therefore, $R_0 = \{0, 1, 2, ….t\}$ and $R_1 = \{t+1, t+2 … L-1\}$. The within-class variance, between class variance, and the total variance are denoted by σ_{2w}, σ_{2b}, and σ_{2T} respectively. So an optimal value of threshold is obtained by minimizing one of the following (equivalent) criterion functions with respect to :

$$t = \text{Arg Min } \eta \qquad \text{……………………………..(4)}$$

The value of η lies between 0 & 1. There are two limits in the value of η. Where when η= upper limit = 1, then only two-valued images are given and when η= lower limit = 0, then only single constant gray-level is given.

5 Hierarchical Self Organizing Mapping (HSOM)

A self organizing map (SOM) comes under unsupervised learning of feedback networks. SOM is another type of neural networks also known as SOFM (Self Organizing Feature Maps).As in the brain, SOM also has self organization property. There is direct connection between input and output devices, and output nodes are also interconnected (different from general feed forward NN).The weights of the output nodes will be adjusted based on the input connected to them, and also the weights of the neighborhood output nodes. Therefore, output nodes will be ordered in a natural manner. Similar nodes will be close to each other. There are two modes in which SOMs operate, such as : training and mapping. Training is a competitive process, also called vector quantization. Mapping automatically classifies a new input vector. The HSOM is the extension of the conventional self organizing map used to classify the image row by row. In this lowest level of weight vector, a higher value of tumor pixels, computation speed is achieved by the HSOM with vector quantization. In the field of medical, segmentation has wide application. The hierarchical self organizing map has been used for multi scale image segmentation. The combination of self organization and graphic mapping technique is known as HSOM. Here hybrid technique is used which has the advantages of HSOM, so as to implement for the MRI image segmentation. MR brain image is loaded into MATLAB 7.0. in the form of matrix. Next initialize the variables sigma, weight vector and winning neuron .In that Calculate the neighborhood function, weight vector and winning neuron .Here neuron is the input and winning neuron is the output [3].
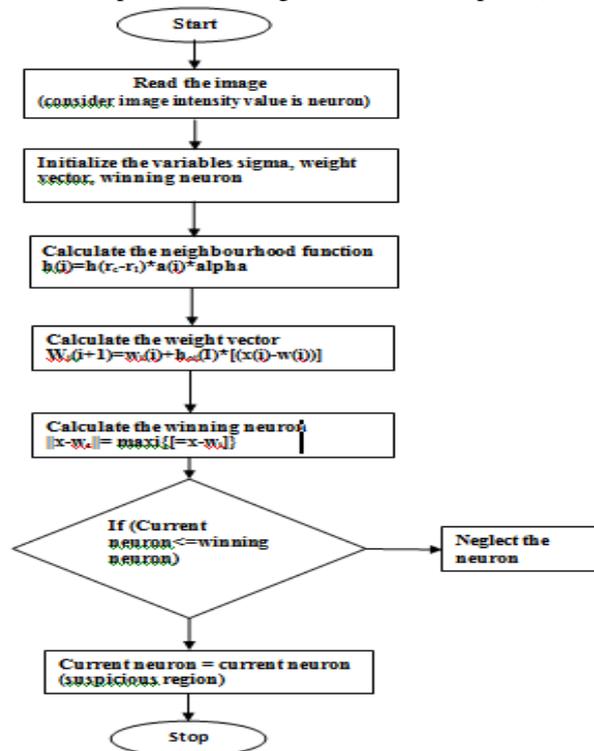
Fig 3.1 : Flow chart of HSOM Method

IV.   Results And Discussion

The proposed algorithms have been implemented using MATLAB. Here various methods of image segmentation are analyzed and discussed. In this system, based on the power of Artificial Neural Network (ANN), a computer-aided brain tumor diagnosis based on hierarchical Self Organizing Mapping is done. It is too much complex to measure the segmentation of image because of no common algorithm. So in order to measure the quality of segmentation of image the statistical measurements are used. The experiment is conducted over the five MRI images using the algorithms Hierarchical Self Organizing mapping, Region Growing, Otsu,K Means and FCM. To evaluate the performance various parameter values are found such as rand index (Ri), global consistency error (Gce), segmentation accuracy (Sa), variations of information (Vi), and area (a).

Rand Index : The Rand index (RI) counts the fraction of pairs of pixels and pixels are those whose labeling are consistent between the segmentation which was computed and the ground truth averaging across multiple ground truth segmentations.

Global Consistency Error : The Global consistency error (Gce) measures the extent by which one segmentation can be viewed as a refinement of the other.

Variation of Information : The variation of information (Vi) represents the distance between two segmentations as the average conditional entropy of one segmentation given the other, and as a result measures the amount of randomness in one segmentation which cannot be explained by the other.

In this paper the comparison of methods i.e. cluster based algorithms, thresholding algorithms and hierarchal self organizing mapping algorithms was done and HSOM was found as the best method for image segmentation .Fig 4.1(a) shows the original MRI image, Fig 4.1(b) shows the ground truth image, Fig 4.1(c) shows the segmented image of HSOM method, Fig4.1(d), (e),(f) and (g) shows the segmented images of Region Growing, Otsu,K Means and FCM respectivly.Table4.1 shows the Segmentation Accuracy of different MRI Images which is best for HSOM method as compared to other methods. Table 4.2 shows Areas of Ground Truth Images and Segmented Images.Table 4.3 shows Performance Parameters for different methods.Table 4.4 shows Average values of Ri, Gce and Vi for various methods. The Graph 4.1 shows the comparison of performance parameters for various methods. The Graph shows the comparison of average values of rand index, average values of global consistency error,and average values of variation of information for various methods. For better performance Ri should be higher comparatively, Gce and Vi should be lower which is clear from the Graph 4.1. The performance parameters chart reveals that the rand index of hierarchical self organizing mapping is higher than others while the global consistency error and variation of information error of hierarchical self organizing mapping is lower than others.

Fig 4.1: Results of Different MRI Images by applying various methods shown in the fig (a) Original Image, fig (b) Ground Truth image fig (c) to fig (g) Segmented Images of HSOM, Region Growing , Otsu , K Means , FCM

For image-1

Fig (a)   (b)   (c)   (d)   (e)   (f)   (g)
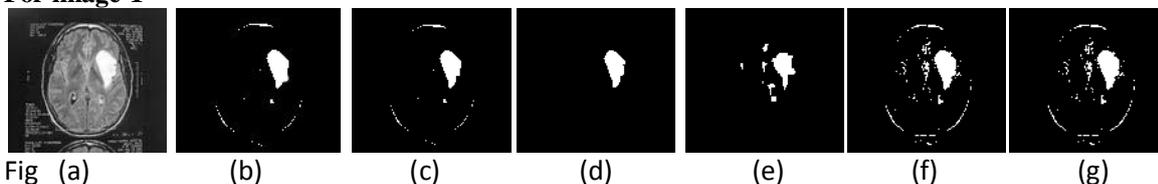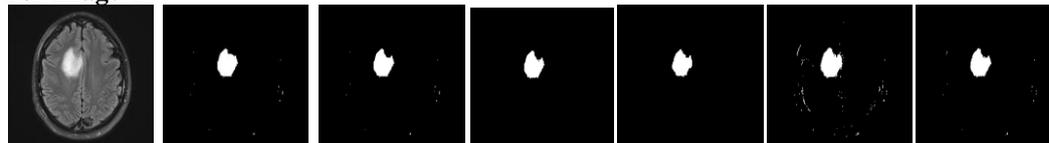
For image-2

For image-3
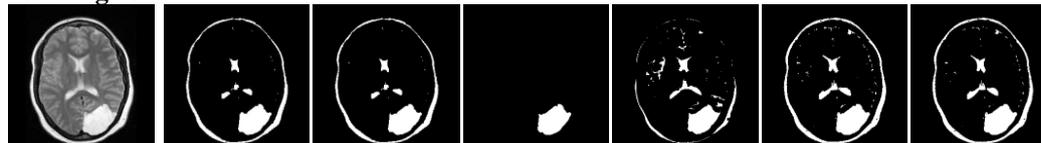
**For image-4**
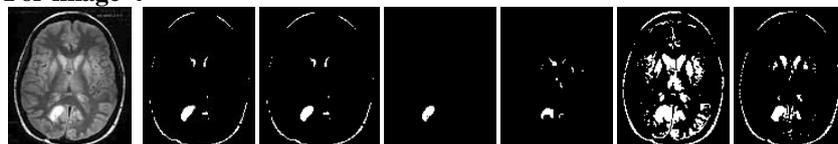
For image-5
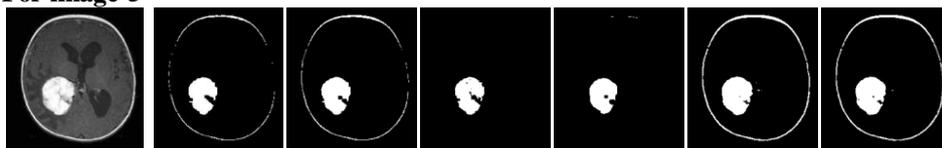
Table-4.1 : Segmentation Accuracy of different MRI Images

Images	Methods	Hierarchical Self Organizing Mapping	Region Growing	Otsu	K means	Fuzzy C means
1	Segmentation Accuracy	94.6622	91.7153	89.8599	167.5377	107.9257
2		96.2918	96.0347	91.2184	93.5783	97.9332
3		99.0744	93.5562	89.6722	70.1631	103.721
4		93.6552	89.3195	67.9729	169.8229	135.7683
5		99.4895	100.0756	96.8309	70.7531	116.6366

Table-4.2 : Areas of Ground Truth and Segmented Images

| Images	Methods	Hierarchical Self Organizing Mapping	Region Growing	Otsu	K means	Fuzzy C means
1	Area 1 (Ground Truth Image)	8.9322	8.9415	8.9608	8.948	8.944
	Area 2	8.4554	8.2007	8.0522	14.9913	9.6529
2	Area 1 (Ground Truth Image)	34.2238	34.1763	34.1872	34.2281	34.2142
	Area 2	32.9548	32.8242	31.185	32.0301	33.5071
3	Area 1 (Ground Truth Image)	82.8334	83.0213	83.3506	83.3506	83.3621
	Area 2	82.0667	77.6715	74.7423	58.4813	86.464
4	Area 1 (Ground Truth Image)	2.9107	2.9229	2.913	2.9107	2.9129
	Area 2	2.726	2.6107	1.98	4.93	3.9548
5	Area 1 (Ground Truth Image)	54.7087	54.7087	54.6859	54.6933	54.6814
	Area 2	54.4274	54.75	52.93	38.6972	63.7785

Table 4.3 : Performance Parameters for different methods

Images	Methods/ Parameters	Hierarchical Self Organizing Mapping	Region Growing	Otsu	K means	Fuzzy C means
1	Rand Index	0.9037	0.7866	0.7118	0.7438	0.7499
	Global Consistency Error	0.2713	0.4426	0.5639	0.5305	0.5458
	Variation of Information	1.7587	3.3427	4.4336	4.3895	4.5025
2	Rand Index	0.956	0.8808	0.7845	0.7767	0.834
	Global Consistency Error	0.1365	0.2957	0.3957	0.4105	0.3457
	Variation of Information	0.9222	2.2362	3.2923	3.7277	2.8733
3	Rand Index	0.9764	0.8009	0.7687	0.818	0.8347
	Global Consistency Error	0.1134	0.3446	0.453	0.3919	0.3738
	Variation of Information	0.7604	3.1271	3.8279	3.3293	3.0942
4	Rand Index	0.9228	0.7184	0.6667	0.6793	0.7108
	Global Consistency Error	0.1543	0.3534	0.4765	0.6056	0.5579
	Variation of Information	1.1188	3.0827	3.9671	4.9737	4.1204
5	Rand Index	0.9794	0.823	0.7865	0.7735	0.776
	Global Consistency Error	0.1102	0.3734	0.3951	0.3691	0.3576
	Variation of Information	0.6893	3.1773	3.4585	3.3129	3.2637

Table 4.4: Average values of Parameters for different methods

Methods	Rand Index	Global Consistency Error	Variation of Information
Hierarchical Self Organizing Mapping	0.94766	0.15714	1.0498
Region Growing	0.80194	0.36194	2.9932
OTSU	0.74364	0.45684	3.79714
K MEANS	0.75826	0.46152	3.94662
FUZZY CMEANS	0.78108	0.43616	4.5025

Graph 4.1 : Comparison of Average Values of Ri, Gce & Vi from Various methods
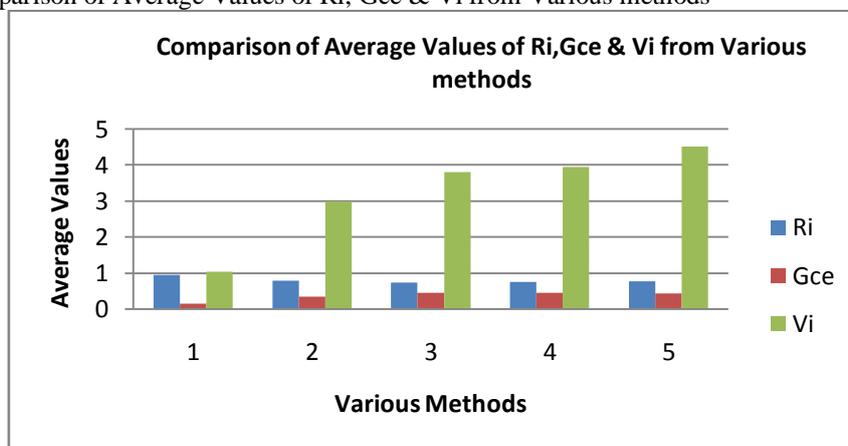
In this graph 4.1 axis of various methods showing 1, 2 ,3 ,4 and 5 which represents HSOM, Region Growing , Otsu, K Means and Fuzzy CMeans methods respectively.

References

[1] Arjan Simonetti , "Investigation of brain tumor classification and its reliability using chemometrics on MR spectroscopy and MR imaging data",2004.

[2] Noworolski , S.M.; Nelson, S.J.; Henry, R.G.; Day, M.R.; Wald, L.L.; Star-Lack, J.; Vigneron, D.B. Magnetic resonance in medicine 1999, 41, 21-29.

[3] T.Logeswari1 and M.Karnan2,"An improved implementation of brain tumor detection using segmentation based on soft computing",Journal of Cancer Research and Experimental Oncology Vol. 2 pp. 06-014, March, 2010.

[4] B.Sathya and R. Manavalan ,"Image Segmentation by Clustering Methods: Performance Analysis", IJCA vol 29-No.11,September 2011.

[5] Rafael C.Gonzalez and Richard E. Woods,"Digital Image Processing",Second Edition,Prentice Hall of India Private Limited ,New Delhi – 2007.

[6] OtSu NA. "Threshold selection method from aray level histogram". IEEE Trans. On system, man, and cybernetics. Vol. 9. No. 1. Pp. 62-66, 1979.

[7] Krishna Kant Singh and Akansha Singh, "A study of Image Segmentation Algorithms for Different Types Of Images", IJCSI International Journal of Computer Science Issues 2010.

[8] Mrs. Bharati R. JipKate et al., A comparative Analysis of Fuzzy C-Means Clustering and K-Means Clustering Algorithm", IJCER / May-June, 2012 / Vol. 21 / Issue No. 3 / 737-739.

[9] M. Sezgin and B. Sankur (2004). "Survey over image thresholding techniques and quantitative performance evaluation". Journal of Electronic Imaging 13(1) : 146-165.