Research Article Impact Factor: 4.226 ISSN: 2319-507X
Abdul Kalam Abdul Salam, IJPRET, 2015; Volume 3 (9): 1690-1702 IJPRET

INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

COMPARATIVE STUDY OF BRAIN TUMOR DETECTION USINGNON-MEDICAL IMAGES AND DICOM IMAGES

ABDUL KALAM ABDUL SALAM[1], ANIL V. DEORANKAR[2], JAYKUMAR LACHURE[3]

Accepted Date: 05/03/2015; Published Date: 01/05/2015

Abstract: Image processing has become an area of vast possibilities to explore as the advances in research field in this domain are gaining momentum. Brain tumour detection is a critical task these days. We all know that image processing can be applied on both non-Medical Images (jpeg, png, bnp, etc) as well as Medical Images(such as DICOM).This paper focuses on the comparative study of algorithms K means, Fuzzy C means and Hierarchical clustering on Non-Medical as well as DICOM Images. Finally the tumour area is specified as confirmation step. A user friendly MATLAB GUI program has been constructed to test the proposed algorithm.

Keywords: Brain tumour, DICOM, Non-Medical Images, Rough Set Theory.

Corresponding Author: MR. ABDUL KALAM ABDUL SALAM

Access Online On:

www.ijpret.com

How to Cite This Article:

Abdul Kalam Abdul Salam, IJPRET, 2015; Volume 3 (9): 1690-1702

PAPER-QR CODE

INTRODUCTION

The brain is anintricate organ as it contains 50-100 billion neurons forming a gigantic neural network. Detection of anatomical brain structures with their precise location is important for treatments like radiation therapy and surgery. Radiologists do the diagnosis of brain tumour manually on MRI images but it being time consuming and error prone as huge number of image slices and the big variations between them. DICOM (Digital Imaging and Communications in Medicine) plays an important role as it is the standard for handling it, storing it, reading it, viewing it and writing it, printing information for medical Imaging. The techniques like MRI (Magnetic Resonance Imaging), NMRI (Nuclear Magnetic Resonance Imaging), MRT (Magnetic Resonance Tomography) and CT (Computed Tomography) Scan are being broadly used to get the images for processing to identify the tumour, out of which MRI is commonly used as it provides much greater contrast between the diverse soft tissues of the body compared to computed tomography (CT). Segmentation techniques like K means clustering, Fuzzy C means, Hierarchical, Watershed Algorithms, and Self Organizing Maps are widely implemented depending on which methodology is required as it can be Region growing, frequency domain based or thresholding based which classifies the MRI Images.

A tumour is an acronym for a neoplasm or a solid lesion formed by an abnormal growth of cells (termed neoplastic) which looks like a swelling. Brain Tumours are composed of the cells that exhibit abnormal and unrestrained cell division. Brain tumour can be benign or malignant, benign being non-cancerous and malignant are cancerous. Malignant tumours are classified in to two types like Primary and Secondary tumours. Benign tumour is less harmful compared to malignant as in malignant tumour it spreads rapidly invading other tissues of brain, progressively aggravating the condition causing death.

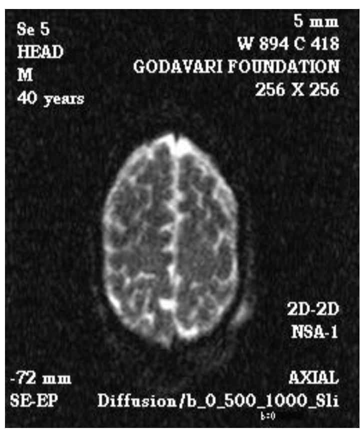
MRI images to detect brain tumour classifies the tumourimplying on whether the brain is an abnormal tissue having normal volume brain tissues like white matter, gray matter and CSF (cerebrospinal fluid) but also have some slices hold pathology like edema and necrosis thus making them abnormal brain tissues. On the basis of CSF Symmetry on the vertical axis through the brain center a normal volume brain tissue and an abnormal volume brain tissue could be classified. The MRI images can be of T1, T2 weighted type of which T2 weighted Images are being extensively used in Medical Imaging as in the case of cerebral and spinal study, the CSF (cerebrospinal fluid) are lighter in T2 weighted images as they are acquired using fast echo spin sequence whereas the T1 weighted images are acquired using a spin echo sequence. Primary focus on exact brain tumours location and its extraction with parameters like area and time to yield faithful and error free output.

Segmentation is a process of dividing the image into different parts having alike features. The pre-processing stages needs to be done on the image at first, followed by clustering algorithms and towards the fag end thresholding be done for the extraction of the tumour which is the region of interest (ROI) from the complete image. The features for thresholding being intensity based, area based. Thresholding is very important part of segmentation as the tumour must be isolated from the brain image.

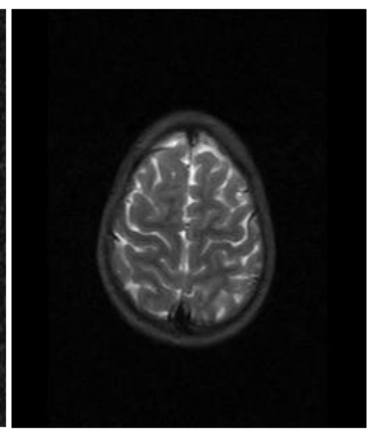
The below is a flowchart (sequence) of precisely how a tumour is detected in ours system.

2. PRE PROCESSING:

The MRI image contain labels on the MRI such as patient name, age and marks and some other information which could get in the way with the tumour detection is not of interest when detecting a tumour. At this time Pre-processing is of imperative importance as the intensity value, larger than that of the threshold value is detached from MRI right from the first row and column of the image. The high intensity values of film artifact are removed from MRI brain image.



i)Original Image



ii)Pre-Processed Image

Tumour location detection algorithm:

Tumour location detection algorithm depends on the fact that human brain is symmetric about vertical axis. In the given system, this algorithm is exploited before applying segmentation methods. If this algorithm gives location of tumour then only apply segmentation methods.

Step 1- Given image divided in four equal parts.

Step 2- As human brain is symmetric about vertical axis, upper two parts compared on basis of number of pixels present in each intensity level i.e. histogram matching.

Step 3- Step 2 repeated for lower two parts.

Step 4- if mismatch found in upper two parts' histogram comparison, symmetry disturbed in upper part hence 'tumour present in upper part'

Else if mismatch found in lower two parts' histogram comparison, 'tumour present in lower part'

Else 'tumour not present' in given image.

3. SEGMENTATION TECHNIQUES

Cluster analysis or clustering is the charge of assigning a set of objects into groups (called clusters), so that the objects in the same cluster are more similar to each other than to those in other clusters. Clustering can be achieved by various algorithms that differ significantly in their notion of what constitutes a cluster and how to efficiently find them. A cluster is therefore a collection of objects which are "similar" between them and are "dissimilar" to the objects belonging to other clusters.

Now for clustering we need to have basis for clustering like criterion function which defines the cluster classification on a stand, keeping in mind that the criterion function should be optimized. Suppose that we have a set D of n samples x1,x2,...xn that we want to partition into exactly into 'c' disjoint subsets D1,..,Dc.Some of the Criterion functions for clustering are as the sum of squared error criterion, Related minimum variance criterion,Scattering criterion.

K means Clustering along with DWT

K means is the clustering method which forms k clusters of n pixel objects, wherein each pixel object belongs to the cluster of the nearest mean, which results of portioning the data space in Voronoi cells. Given a set of observations (x1, x2, ...,xn), where each observation is a d-dimensional real vector. K-means clustering aims to partition the n observations into k sets (k ≤ n) S = S1, S2,...,Sk so as to minimize the within-cluster sum of squares

$$arg_s \min \sum_{i=1}^{k} \sum_{m}^{n} \|X_j - \mu_i\|^2$$

Where $\mu i$is the mean of points in Si. m=$xj\epsilon Si$

It is a simple clustering method and gives fast outputs as well, but the problem of choosing the correct clustered image is a big issue in K means. However if we are using a DWT (Discrete Wavelet Transform) along with k means the we get the high level details of the tumour and also frequency information and time space localization. When a wavelet transform is applied on the MRI image it yields wavelet decomposed image resulting in four sub bands, which are the LL (Lower resolution version of image), LH (Horizontal edge data), HL (Vertical edge data) & HH (Diagonal edge data) sub bands representing approximation, horizontal, vertical and diagonal components in the form of coefficients, respectively. LL sub band contains low level and the other three (LH, HL and HH) contain high level details. DWT (Discrete wavelet Transform) applies the approximation coefficients in LL equal to zero and apply inverse wavelet transform to obtain a high pass image from the remaining (horizontal, vertical and diagonal) sub bands and the resultant image obtained is the level-1 (L1) detail image. Thus DWT (Discrete wavelet Transform) gives a sharpened image which is added along with the original image and to the resulting outcome of the addition segmentation of k means clustering is performed, followed by thresholding techniques to extract the tumour. Thresholding can be based on many methodologies-which can be area based or intensity based.

Algorithm for K means Clustering:

Step 1: Choose K centroids at random from input MR image.

Step 2: Make initial partition of objects into k clusters by assigning objects to closest centroid

Step 3: Calculate the centroid (mean) of each of the k clusters.

a. For object i, calculate its distance to each of the centroids.

b. Allocate object i to cluster with closest centroid.

c. If object was reallocated, recalculate centroid based on new cluster.

Step 4: Repeat 3 for object i = 1,...., N

Step 5: Repeat 3 and 4 until no reallocations occur.

Step 6: Assess cluster structure for fit and stability:

Step 7: Separate Image into K sub images according to clustered indexed Image

Step 8: Apply intensity and area based threshold to extract exact tumour part from image.

Fuzzy C Means

In the year 1973 Dunn developed the Fuzzy C Means algorithm and later in 1981 it was enhanced by Bezdek. However the Fuzzy logic was proposed in 1965 by Lofti A Zadak a professor of Computer Science at University of California, Berkeley.

Fuzzy logic is a form of many-valued logic or probabilistic logic. It by definition only means approximate values rather than fixed and exact. In contrast with traditional logic they can have varying values, where binary sets have two-valued logic, true or false, fuzzy logic variables may have a truth value that ranges in degree between 0 and 1. Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely true and completely false.

Let X be a space of points, with a generic element of X denoted by x. Thus, X=$\{x\}$.A Fuzzy set A in X characterized by a membership function $\mu A(x)$which associates with each point in X a real number in the interval [0,1], with value of $\mu A(x)$ at x representing the grade of membership of x in A. Thus, nearer the value of $(x)$to unity, the higher the grade of membership of x in A. In the hard clustering process, each data sample is assigned to only one cluster and all clusters are regarded as disjoint collection of the data set. In practice there are many cases, in which the clusters are not completely disjoint and the data could be classified as belonging to one cluster almost as well to another.

This algorithm works by assigning membership to each data point corresponding to each cluster centre on the basis of distance between the cluster centre and the data point. More the data is near to the cluster centre more is its membership towards the particular cluster centre. Clearly, summation of membership of each data point should be equal to one. After each iteration, the up-gradation of the membership and cluster centres is done.

Parameters:

n : is the number of data points.

 : represents theclustercentre.

m : is the fuzziness index m ∈[1, ∞].

c : represents the number of cluster centre.

μij: represents the membership ofdatatoclustercentre.

dij: represents the Euclidean distance between ithandjth data and cluster centre.

Main objective of fuzzy c-means algorithm is to minimize:

$$J(U,V)= \sum_{i=1}^{n}\sum_{j=1}^{c}(\mu_{ij})^{m} \|x_i - v_j\|^2$$

Where $\|xi - vj\|$, is the Euclidean distance between ithdata and jthclustercentre.

Algorithmic steps for Fuzzy c-means clustering:

Let X ={x1, x2, x3 ..., xx}be the set of data points and V = {v1, v2, v3 ..., vv}be the set of centres.

Step 1: Randomly select c cluster centers.

Step 2: Calculate the fuzzy membership function μijusing:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{ik}}\right)^{\frac{2}{m-1}}}$$

Step 3: Compute the fuzzy centers vj using:

$$v_j = \left(\sum_{i=1}^{n}(\mu_{ij})^{m} x_i\right) \div \left(\sum_{i=1}^{n}(\mu_{ij})^{m}\right),$$

$\forall j$=1,2,..c.

Step 4: Repeat Step 2&3 until the minimum 'J' value is achieved or $\|Uk+1-Uk\|<\beta$ where, k: is the iteration step.

β : is the termination criterion between [0, 1].

U = $(\mu_{ij})_{n+c}$ is the fuzzy membership matrix.

J: is the objective function.

Hierarchical Clustering:

A Hierarchical clustering method works by grouping data objects into a tree of clusters. There are two types of clustering 1.Agglomerative 2.Divisive. Agglomerative clustering differs in the similarity measures which employ single link, complete link, group average, centroid similarity.

Hierarchical clustering doesn't require specifying the number of clusters. It is deterministic. In agglomerative clustering each element is treated as a singleton cluster and then merged (agglomerated) until all merge in a single cluster, which results in dendograms formation. Dendograms are horizontal lines which when cut at a point you get a specific part or element and explains how clustering helps forming an image.

The Algorithm for Agglomerative Hierarchical Clustering:

Step 1: We first compute the $N \times N$ similarity matrix C.

Step 2: The algorithm then executes $N - 1$ steps of merging the currently most similar clusters.

Step 3: In each iteration, the two most similar clusters are merged and the rows and columns of the merged cluster i in C are updated.

Step 4: The clustering is stored as a list of merges in A.

Step 5: I indicates which clusters are still available to be merged.

Step 6: The function SIM(i,m,j) computes the similarity of cluster j with the merge of clusters i and m.

Step 7: For some HAC algorithms, SIM (i,m, j) is simply a function of $C[j][i]$ and $C[j][m]$, for example, the maximum of these two values for single-link.

Divisive hierarchical is more efficient as we can stop when we reach our goal. Unlike agglomerative clustering we do not need to construct the entire hierarchy all the way down to individual leaves. Also along with efficiency even in accuracy point divisive is better than agglomerative hierarchical clustering.

Algorithm: Divisive Hierarchical Clustering

Step 1: The whole image is in one cluster.

Step 2: Find the most dissimilar point in the image and divide the image into two clusters.

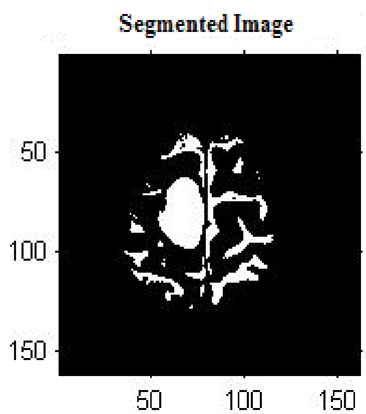
Step 3: Repeat step2 for each cluster.

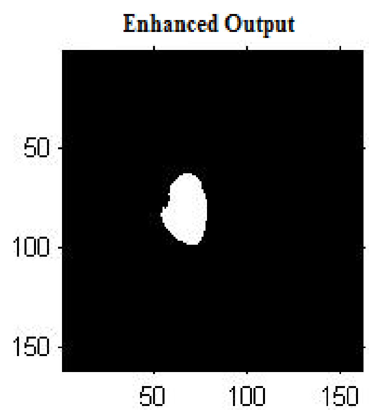
Step 4: A tree like structure is formed. Repeat step 2 until level 4 is reached. Level 4 has 8 clusters.

Step 5: Continue until the tumour cluster is obtained.

During removal of film artifacts, the image consists of noise. The image is given to enhancement stage for the removing high intensity component and the above noise. This part is used to enhance the smoothness towards piecewise homogeneous region and reduce the edge blurring effects, which tells the necessity of Enhancement after segmentation Techniques being applied. Noise is a problem with MRI images also needs to be removed, along with the quality of the image to be maintained. Hence Enhancement step is needed.

When the tumour is detected, because of intensity based techniques are used, small matters in brain MRI like soft tissues as well as big objects like eye remain in the image alongside the tumour. In our system, area based thresholding is used in post processing of tumour image to





4. RESULTS:

The Clustering algorithms like K means, Fuzzy C means and hierarchical clustering were applied on the database of 60 brain tumour images in non-medical format (.jpg, .png, .bmp etc.) as well as Images in DICOM format yielding the following efficiencies on the basis of four parameters described below:

TP: True Positive: Tumour Present and detected

TN: True Negative: Tumour not present and not detected

FP: False Positive: Tumour not present and detected

FN: False Positive: Tumour present and not detected

Efficiency= (TN+TP)/(TP+TN+FP+FN)

Table 1: Efficiency in Non-medical format Images

Method	Efficiency
K means	70%
Fuzzy C means	70%
Hierarchical Clustering	66.66%

Table 2: Efficiency in DICOM Images

Method	Efficiency
K means	86.66%
Fuzzy C means	83.33%
Hierarchical Clustering	83.33%

The timing parameter for an image from the database is as

Table 3: Average Time required

Method	Time in Seconds
K-Means	0.8608
Fuzzy C Means	1.6363
Hierarchical Clustering	0.4068

It clearly depicts that the time required for Hierarchical is least and that for Fuzzy C means is maximum to detect the tumour. But tumour detection is much more prominent in FCM than other methods.

Algorithm	Advantage	Disadvantage
K-Means	Simple and Less Computation is required. Output is almost accurate every time	Efficiency is less in noisy MRI images.
Fuzzy C Means	Utilizes the advantages of Fuzzy Set over the Crisp Set. Most Accurate detection of edges of tumour	Produce better result than K-means but not Robust to noisy images.
Hierarchical	Computational time is much less and hence fastest as compared to all	Output of tumour is not exact in many cases reducing efficiency.

5. CONCLUSION

A novel system that can be used as a second decision for the surgeons and radiologists is proposed. In this system brain tumours have been segmented with the help of three methods. The execution time for Hierarchical clustering was less compared to the other clustering methods. Regarding the number of tumour pixels, K-means clustering and Fuzzy C means gave a better result than the other methods. The three clustering algorithms were tested with a database of MRI brain images. For non-medical format images (.jpg, .png, .bmp, etc.), K-means and Fuzzy C means clustering achieved about 70% result. Hierarchical clustering achieved about 66.66% and for DICOM images overall efficiency for all algorithms proved to be improved up to 86%.Tumour location detection algorithm increases overall efficiency of the system as it reduces FP(False Positive) output form segmentation algorithm. The texture can be taken as an additional parameter for tumour detection**.** It clearly depicts that the time required for Hierarchical is least and that for Fuzzy C means is maximum to detect the tumour. But tumour detection is much more prominent in FCM than other methods. When we move from non-medical images to medical images i.e. DICOM, the result is more promising.

REFERENCES

1. T. Rajesh, R. Suja Mani Malar "Rough Set Theory and Feed Forward Neural Network Based Brain Tumor Detection in Magnetic Resonance Images" Proceedings of the International

Conference on Advanced Nanomaterials& Emerging Engineering Technologies" (ICANMEET-2013).

2. X. D. Yue, D.Q.Miao, N.Zhang "Multiscale roughness measure for color image segmentation" International Journal of Information Science(ELSEVIER) 2012

3. KshitijBhagwat, Dhanshri More, SayaliShinde, AkshayDaga, Assistant Prof. RupaliTornekar , " Comparative Study of Brain Tumor Detection Using K Means , Fuzzy C Means and Hierarchical Clustering Algorithms," *International Journal of Scientific & Engineering Research , Volume 2,Issue 6, pp. 626-632 , June 2013.*

4. S. Roy and S. K. Bandoyopadhyay, " Detection and qualification of Brain Tumor from MRI of Brain and Symmetric Analysis" International Journal of Information and communication Technology Research, Volume 2 No.6, June 2012

5. Aboul Ella Hassanien,AjithAbraham,"Rough Set and Near sets in Medical Imaging: A Review",IEEE Trans. On Information Technology in Biomedice,2009

6. Monika Sinha, KhushbooMathur , " Improved Brain Tumor Detection With Ontology," *International Journal Of Computational Engineering Research ISSN 2250–3005 Vol. 2, Issue No.2, pp. 584-588, Mar-April-2012*

7. Arati Kothari, "Detection and Classification of Brain Cancer using Artificial Neural Network in MRI Images," *World Journal of Science and Technology , ISSN 2231-2587, pp. 1-4, 2012*

8. M. Gopu, T. Rajesh,"Brain Tumor Segmentation based on Rough Set Theory for MR Images with CA Approach," *International Journal of Emerging Trends in Electrical and Electronics, vol. 4, issue. 1, pp. 71-76, June 2013.*

9. E. Venkateshwara Reddy and Dr. E.S. Reddy, " Image Segmentation using Rough Set based Fuzzy K Means Algorithm," *Global Journal of computer science and technology ,volume 13,Issue 6 ,version 1.0, pp. 23-29, 2013*

10. A. Jayachandran, R. Dhanasekaran, " Brain Tumor Detection and Classification of MR Images Using Texture Features and Fuzzy SVM classifier," *Research Journal of Applied Sciences, Engineering and Technology,ISSN 2040-7459, pp. 2264-2269 , January 2013.*

11. J. Luts, T. Laudadio, A. J. Idema, A. W. Simonetti, A. Heerschap, D. Vandermeulen and S. Huffel, "Nosologic Imaging of the Brain Segmentation and Classification using MRI and MRSI," *Journal of NMR in Biomedicine, vol. 22, issue 4, pp. 374-390, May 2009.*

12. Ahmed Kharrat, KarimGasmi, "A Hybrid Approach for Automatic Classification of Brain MRI Using Genetic Algorithm and Support Vector Machine," *Journal of Sciences, pp.71-82, 2010.*

13. P. Tamijeselvy, "Performance Analysis of Clustring Algorithms in Brain Tumor Detection of MR Images," *European Journal of scientific research , ISSN 1450-216X,vol. 62, no. 3, pp. 321-330, 2011*

14. Z. Shi, L. He, T. N. K Suzuki, and H. Ito, "Survey on Neural Networks used for Medical Image Processing," *International Journal of Computational Science, 2009.*