Detection of Brain Tumor using Modified K Means Algorithm and SVM

Labeed K Abdulgafoor
M tech DSP 4th semester,
KMCT college of engineering
Calicut. Kerala

Aji George
Assistant professor, department of ECE
KMCT College of engineering
Calicut, Kerala

ABSTRACT
This paper presents a research on MRI images using wavelet transformation and modified K – means clustering for tumor segmentation. The first step is to perform image segmentation. It allows distinguishing masses and micro calcifications from background tissue. In this paper wavelet transformation and K-means clustering algorithm have been used for intensity based segmentation. The proposed algorithm is robust against noise. In this case, discrete wavelet transform (DWT) is used to extract high level details from MRI images. The processed image is added to the original image to get the sharpened image. Then modified K-means algorithm is applied to the sharpened image in which the tumor region can be located.. The combination of noise-robust nature of applied processes and the modified K-means algorithm, and SVM gives better results.

Keywords
support vector mechine(svm),discrete wavelet transform(DWT),modified K means algorithm

1. INTRODUCTION
The influence and impact of digital images on modem society is tremendous, and image processing is now a critical component in science and technology. Several new complex medical imaging modalities, such as X-ray, magnetic resonance imaging (MRl), and ultrasound, strongly depend on computer technology to generate or display digital images. With computer techniques, multidimensional digital images of physiological structures can be processed and manipulated to help visualize hidden diagnostic features that are otherwise difficult or impossible to identify using planar imaging methods. Image segmentation may be defined as a technique, which partitions a given image into a finite number of non- overlapping regions with respect to some characteristics, such as gray value distribution, texture. segmentation of medical images is required for many medical diagnoses like radiation treatment, planning volume visualization of regions of interest (ROI) defining boundary of brain tumor and intra cerebral brain hemorrhage, etc. Basically, image segmentation methods can be classified into three categories: edge-based methods, region based methods and pixel-based methods. Modified K-means clustering is a key technique in pixel-based methods. These are simple and the computational complexity is relatively low compared with other region-based or edge-based methods, the application is more practicable. Furthermore, K-means clustering is suitable for biomedical image segmentation as the number of clusters is usually known for images of particular regions of the human anatomy. It is an unsupervised clustering algorithm that classifies the input data points into multiple classes based on their inherent distance from each other.

In this paper a hybrid technique is used that makes use of DWT and modified K-means algorithm. By using DWT we extract the high pass image and then this image is applied as the input of modified K-means algorithm for segmentation. Our proposed method utilizes the advantage of noise-robust nature of wavelet and the simplicity of K-means algorithm which results in better detection of tumor in MRI images

2. PROPOSED WORK

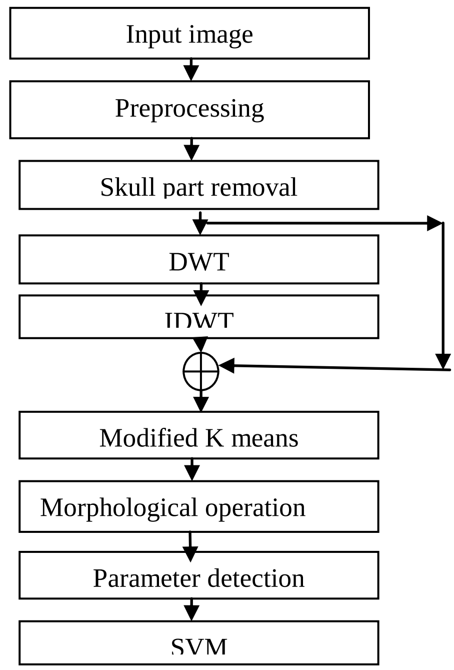


Fig .1 ***flow chart of proposed work***

2.1. Image preprocessing

This step involves apply the pre processing filters like mean, median Gabor filters to increase the clarity of image [2],and reduce the unwanted pixels from the image.

2.2. Skull Removal
This is pre processing step which is required to produce better results. Skull is outer part of the brain surrounding it i.e. the removal of its non-cerebral tissues[8]. The main problem in skull- stripping is the segmentation of the non-cerebral and the intracranial- tissues due to their homogeneity intensities. So it may affect the result of seed point selection .Some observations

are required to find the range of gray value of skull portion. Following of the steps which are involved in skull removal process:

- First of all find the size of the image and store the no of rows and columns in separate variables.
- Perform iteration for half of the columns and all rows
- Process half of image to convert white pixels into the black pixels by setting their gray value to zero.
- Same steps are repeated for the remaining column and row.

2.3. Wavelet transform

Here we used discrete wavelet transform (DWT)and inverse discrete wavelet transform(IDWT).the wavelet transform can be used for preprocessing also. It is used for to sharpen the image. This transform also improve the speed of the system. The wavelet transform done the following steps [5]

- Wavelet transform is applied to an input MR image to obtain wavelet decomposed image resulting in four sub bands. These are the LL (lower resolution version of image), LH (horizontal edge data), HL(vertical edge data) and 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, HH) contain high level details.
- Set 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. We call the resultant image level-1 (L1) detail image.
- Add L1 to the original image to get a sharpened image.

2.4. Modified K means algorithm

currently the amounts of data stored in databases (online and offline) are so huge that create a crucial need for effective and speedy data analysis methods. Cluster analysis is one of the primary data analysis tasks that helps in interpretation and understanding of natural grouping or structure in a dataset. K means clustering is the most widely used and studied method among clustering formulations that are based on minimizing a formal objective function [4]. Modifications to K means clustering method that makes it faster and more efficient are proposed. The main argument of the proposed modifications is on the reduction of intensive distance computation that takes place at each run (iteration) of K-means algorithm between each data point and all cluster centers [6]. To reduce the intensive distance computation, a simple mechanism by which, at each iteration, the distance between each data point and the cluster nearest to it is computed and recorded in a data structure is suggested. Thus, on the following iteration(s) the distance between each data point and its previous nearest cluster is recomputed.

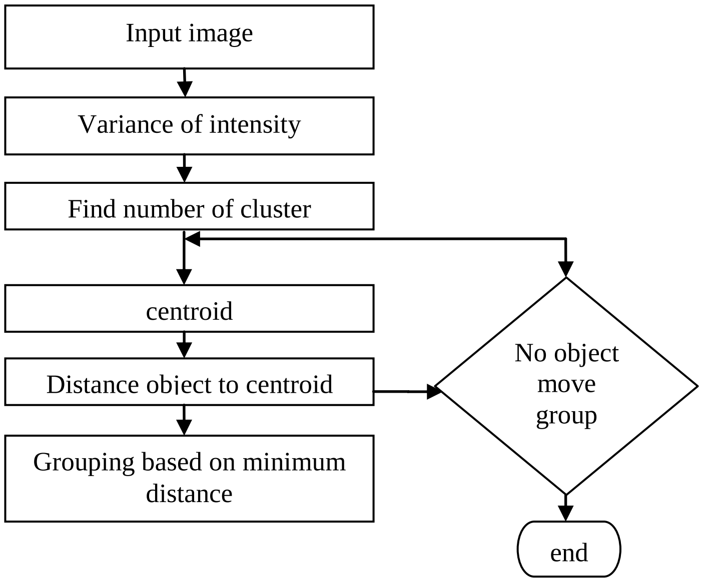


Fig.2. ***flow chart for modified K means algorithm***

2.5. Morphological processing

Morphology mainly deals with the contour and structure of the object. So this is used to perform object extraction, noise removal procedure etc[3] . For the same purpose we are applying these operations to enhance the object boundary and to remove the noise from the image .

The most basic morphological operations are dilation and erosion. Dilation adds pixels to the boundaries of objects in an image, while erosion re-moves pixels on object boundaries. The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. In the morphological dilation and erosion operations, the state of any given pixel in the output image is determined by applying a rule to the corresponding pixel and its neighbors in the input image. The rule used to process the pixels defines the operation as dilation or erosion. One important part in morphological operation is to choose the structuring element. A structuring element is a matrix consisting of only 0's and 1's that can have any arbitrary shape and size. The pixels with values of 1 define the neighborhood .Two-dimensional, or flat, structuring elements are typically much smaller than the image being processed. The centre pixel of the structuring element, called the origin, identifies the pixel of interest -- the pixel being processed. The pixels in the structuring element containing 1's define the neighborhood of the structuring element. In our project work we are taking DISK shape as structuring element. In the operation of image dilation and erosion we are considering disk structuring element of varying radii so that they obtained image is free from small unwanted parts. In MATLAB working environment there are two built in functions used for dilation and erosion. These morphological functions position the origin of structuring element, its center element over the pixel of interest in the input image. For pixels at the border of the image, parts of the neighborhood defined by the structuring element can extend past the border of the image. To process border pixels, the morphological functions assign a value to these undefined pixels, as if the functions had padded the image with additional rows and columns. The value of these padding pixels varies for dilation and erosion operations.

2.6. Parameter detection

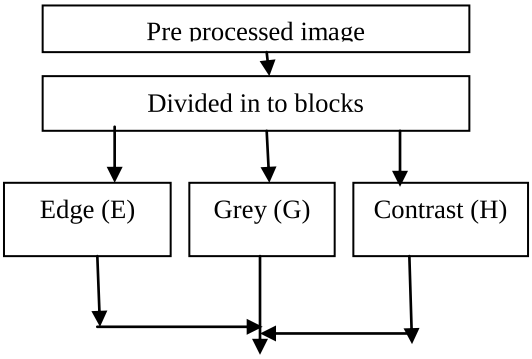


Fig.3. ***parameter detection flow chart***

Recent advances in medical image analysis often include processes for an image to be segmented in terms of a few parameters and into smaller sizes or regions, to address the different aspects of analyzing images into anatomically and pathologically meaningful regions [7]. Classifying regions using their multiparameter values makes the study of the regions of physiological and pathological interest easier and more definable. Here, multiparameter features refer to the following three specific values for the edges (*E*), gray values (*G*), and local contrast (*H*) of the pixels.

2.6.1Edge (E) Parameter

Edge information is often used to determine the boundaries of an object. This is mainly used for analysis to derive similarity criterion for a pre-determined object. In our earlier findings, we observed that the edge pixel for a tumor image is reduced in comparison with that of a non-tumor image. The reduction of edge count for tumor images is because the brain substance has been pushed aside and compressed by the growth of the tumor within a confined space, the intracranial cavity. we use the Sobel edge[1] detection method to detect image edges (*IE*). Image *I1* and *I2* is obtained by filtering an input image with two convolution kernels concomitantly, one to detect changes in vertical contrast (*hx*) and the other to detect horizontal contrast (*hy*), Image output (*IE*) is obtained by calculating the gradient magnitude of each pixel, Subsequently, the edge parameter (*E*) is calculated, whereby *E(r, c)* is increased by one each time when *IE (x, y)* = ‘1’ in a supervised block

2.6.2. Gray (G) Parameter:

The gray parameter (G) for each block of the brain is accumulated, and controlled by a binary image (IT) using the GD value as a threshold,. GD value is calculated using the average pixel value (Iavs) of each image slice (S) for total image slices (T) of an image dataset,

The pixels intensity for each slice was calculated to establish the threshold values and thus provide the basis for analysis of clinical MR images from patients with brain tumors.

2.6.3 Contrast (H) Parameter:

Contrast (H) is often used to characterize the extent of variation in pixel intensity. In the present technique, the computational program analyses the differences, especially in instances of strong dissimilarity, between entities or objects in an image I (x,y).We adopt the minimum/maximum stretch algorithm for the 8-neighborhood connectivity, where min H and max H represent the minimum and maximum intensity values of the neighborhood pixel C8(IH),. In the previous studies, tumor cells are often associated with higher value of contrast (H) parameter Hd is obtained by totaling the contrast of a supervised block.

2.6.4 Classification using support vector machine (SVM)

The aim of classification is to group items that have similar feature values into groups. Classifier achieves this by making a classification decision based on the value of the linear combination of the features [9]. SVM is a binary classification method that takes as input labeled data from two classes and outputs a model file for classifying new unlabeled/labeled data into one of two classes. The SVM originated from the idea of the structural risk minimization that was developed by Vapnik. Support vector machines are primarily two class classifiers that have been shown to be attractive and more systematic to learning linear or non-linear class boundaries. The use of SVM, like any other machine learning technique, involves two basic steps namely training and testing.

Training an SVM Involves feeding known data to the SVM along with previously known decision values, thus forming a finite training set. It is from the training set that an SVM gets its intelligence to classify unknown data. In SVM, for two class classification problem, input data is mapped into higher dimensional space using RBF kernel. Then a hyper plane linear classifier is applied in this transformed space utilizing those patterns vectors that are closest to the decision boundary

3. RESULTS AND DISCUSSION

The tumor affected cells are found out by applying modified K means algorithm using MATLAB simulator.fig 4 shows the MRI of brain tumor.the given image is preprocessed and given to DWT for removing low frequency component as shown in fig 5.the skull part removed image is shown in fig6.the output of modified K means algorithm and total output is shown in fig7 and fig 8.

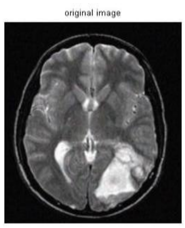


Fig.4. input image **Fig.5. IDWT output**

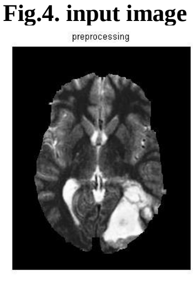
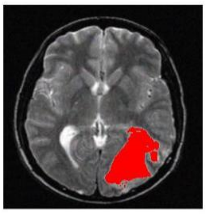


Fig.6. skull pert removed output **Fig.7. Modified K means**



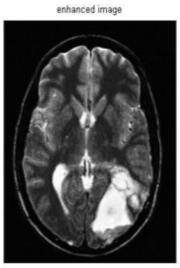


Fig.8. total output **FIG.9. Enhanced**

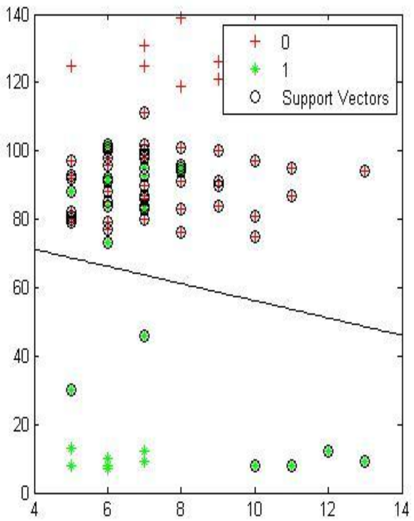


Fig.9.support vectors

4. CONCLUSION

This paper reported a methodology of segmentation of MRI images using wavelet and modified K-means algorithm. Wavelet transform made the algorithm noise free because wavelets provide frequency information as well as time-space localization. In addition, their multi-resolution character enables us to visualize image at various scales and orientations. Resolution reduction using wavelet depends on the amount of noise as well as the area of the target. Then k-means was applied to segment the MRI. K-means provides a very simple and efficient method of segmentation. Thereafter a parameter detection method has been employed to detect the tumor region. We proved our result is better by comparing with other two methods. On the other hand, this paper has shown that advanced technique of image processing and micro calcification detection which is useful in computer aided diagnosis. The intelligent systems development combined with health specialists' knowledge improve diagnostics associated to different pathologies. This method can be easily extended for brain tumor segmentation. In future we may seek to employ this method on SAR images and may improve this algorithm accordingly

5. ACKNOWLEDGEMENT

We would like to appreciate to all of instructors and friends at Department of electronics and communication, professors and assistant professors, head of department, KMCT College of engineering, Calicut, Kerala.

REFERENCES

[1] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Pearson Education, 2002..

[2] Swati Tiwari, Ashish Bansal,Rupali Sagar. Identification of brain tumor using seeded region growing method. Journal of education.volume2, issue 1 Nov 2012

[3] Parisa Beham,Gurulakshmi. Morphological image processing approach on detection of tumour and cancer cells (2011) *IEEE Trans, Image processing*

[4] Ahmed, M. and D.B. Mohammad, 2008. Segmentation of Brain MR Images for Tumor Extraction by Combining K-means Clustering and Perona Malik Anisotropic Diffusion Model. International Journal of Image Processing, 2: 27-34.

[5] Mostafa, M.G., M.F. Tolba, T.F. Gharib and M.A. Mageed, Medical Image Segmentation Using Wavelet Based Mutiresolution EM Algorithm.

[6] Hartigan, J.A., Wong, M.A.: Algorithm AS136: A K-Means Clustering Algorithm. Applied Statistics 28, pp. 100–108 (1979).

[7] Abbasi S. and Mokhtarian F. (2001) *IEEE Trans, Image processing*, 10(1), 131-139.

[8] Abdel-Halim Elamy, Maidong Hu. (2007) *IEEE Image Processing Society*, 872-875.

[9] Bailet J.W., Abemayor E., Jabour B.A., Hawkins R.A., Hoh C.K., Ward P.H. (1992) *Laryngosope*, 102, 281- 288.