Clustering Techniques for Brain Tumor Detection

Ramish B. Kawadiwale¹ and Milind E. Rane²

¹ Vishwakarma Institute of Technology/Dept. of Electronics and Telecommunication, Pune, India
Email: ramish.kawadiwale.vit@gmail.com
² Vishwakarma Institute of Technology/Dept. of Electronics and Telecommunication, Pune, India
Email: me_rane@yahoo.com

Abstract—Brain tumor is a malformed growth of cells within brain which may be cancerous or non-cancerous. The term ‘malformed’ indicates the existence of tumor. The tumor may be benign or malignant and it needs medical support for further classification. Brain tumor must be detected, diagnosed and evaluated in earliest stage. The medical problems become grave if tumor is detected at the later stage. Out of various technologies available for diagnosis of brain tumor, MRI is the preferred technology which enables the diagnosis and evaluation of brain tumor. The current work presents various clustering techniques that are employed to detect brain tumor. The classification involves classification of images into normal and malformed (if detected the tumor). The algorithm deals with steps such as preprocessing, segmentation, feature extraction and classification of MR brain images. Finally, the confirmatory step is specifying the tumor area by technique called region of interest.

Index Terms—MR Brain images, Clustering, K-means Clustering, Hierarchical, Gray matter (GM), Cerebrospinal fluid (CSF), White matter (WM).

I. INTRODUCTION

BRAIN, the central processing unit of human body, is a soft, delicate and spongy mass of tissues. It is a steady place for signals to enter and being processed. A tumor is a detrimental volume formed by abnormal growth of cells which looks like a swelling. A group of malformed cells that grows inside of the brain or around the brain is called brain tumor. Brain tumors are the tenth most common cause of cancer death. Brain tumors can be classified into two general classes depending on the origin of tumors, their pattern of growth and malignancy. Tumors that are developed from cells within the brain itself are primary brain tumors. Whereas, a secondary or metastatic brain tumors are developed when cancer cells from cancer in another part of the body spread to the brain [1]. According to research statistics, the number of brain tumors patients and number of deaths due to it has increased perhaps as much as 300 over past three decades. The most common primary cancers that spread to the brain are lung, breast, unknown primary melanoma and colon cancer. The structure and function of the brain can be studied noninvasively by doctors and researchers using Magnetic Resonance Imaging (MRI). Magnetic Resonance Imaging (MRI) strongly depends on computer technology to generate or display digital images. MRI systems are very important in medical image analysis. The MR image shows the clear peculiarity between the tissues, bones and fluid, so it makes easy to distinguish the tumor part from the image [5].

In medical imaging, an image is captured, digitized and processed for doing segmentation and for extracting important information. In order to haul out the tumor region efficiently the MRI image should be
preprocessed properly. Preprocessing includes removal of unwanted parts of MR brain images such as film artifacts, Skull part of brain, noise etc. Segmentation is an important process in most medical image analysis [5]. It is very difficult to conduct surgery without exact knowledge of shape, size and location of tumor in brain. Clustering is the appropriate method for biomedical image processing since the approximate or exact number of clusters in images of various organs of body is known. Different clustering algorithms are employed for extraction of tumor region. Some algorithms are automatic while some are manual. In this system, various clustering algorithms are applied on MR brain images. After the clustering process, the cluster containing the tumor is selected as the primary segment. Histogram clustering is applied to eliminate the pixels which are not related to the tumor pixels.

II. LITERATURE REVIEW

Various pre-processing and segmentation techniques are cited in literature for medical image processing [5]. Pre-processing techniques comprise de-noising the image, removal of unwanted parts, enhancement etc. Gaussian filters and bilateral filters give good results in smooth area of image [2, 3 and 4]. But edges and textures get blurred due to these filters. Intermediate gray levels are produced across edges when low pass filter is used to smooth black-and-white images, thus produces blurred images. Many papers given in the literature explains about the tracking algorithm. Weighted median (WM) filters are the extension of median filters, which exploits rank-order information and spatial information of input signal [6].

Segmentation techniques based on thresholding are simple and effective for images with different intensities [7, 8, 9 and 10]. However these methods fail for multichannel images. There were many attempts to correct these problems but there are some factors that can complicate the thresholding operation, for example, busyness of gray levels within the object and its background, inadequate contrast etc [15]. Segmentation by region growing is also effective method but not fully automatic as it requires user interaction [11, 12]. Some fully automatic algorithms are available but they perform poor on natural images. Brain tumor and tissue segmentation in MR images have been always an important area of research. Extraction of good features is fundamental need of successful image segmentation [13]. Edge detection based segmentation using various gradient operators such as Sobel’s operator, Prewitt’s operator is discussed in [14]. Some approaches follow the process sequence as image sharpening and then segmentation algorithms [16].

III. PROPOSED METHODOLOGY- PRE-PROCESSING

Pre-processing incorporates stages as shown in flow chart in Fig. 1.

A. Image Acquisition

MR brain images can be obtained from intra-operative magnetic resonance scanner. There are many such systems available. Among these systems, the 0.5T intra-operative magnetic resonance scanner acquires 256 x 256 x 58(0.86mm, 0.86mm, 2.5 mm) T1 weighted images. The intra-operatively acquired 256 x 256 slice has quality similar to the images acquired with a 1.5T conventional scanner, but the downside of these images is that the slice remains thick (2.5 mm). Even though images do not show hefty distortion, images can suffer
from artifacts due to different factors (surgical instruments, hand movement, radio frequency noise from bipolar coagulation). Nowadays it is possible to acquire images with very limited artifacts during neurosurgical scanning. Brain image obtained by MRI scan is read and stored as 2-D array of pixels in MATLAB v2012b. In this system, images displayed are of size 256 x 256 pixels. A grayscale image has intensity values between 0 and 255, with 255 corresponding to white and 0 to black. For black and white images, where intensity values are integers, the lowest entry corresponds to black, the highest to white. MRI scans are stored in JPEG image format.

B. Film Artifacts Removal

The MR brain image consists of film artifacts or labels such as patient’s name, age and remarks. Tracking algorithm is used to remove film artifacts. The major point to be considered is that for T2 weighted MR image, intensity values of label pixels and tumor region pixels fall in range 200-255 i.e. white color. So there is possibility of losing tumor region pixel if we apply this algorithm to entire image. To deal with this problem, pixels in 4 corners of the MR brain image are tracked and algorithm is applied on those pixels only. To set threshold, given image is raster scanned through each pixel, to analyze the intensity value of the pixels, the all pixels those are in 4 corners of MR image and having intensity values greater than threshold value are removed from MRI. This is general algorithm and can be applied to T1 or T2 weighted MR images. Results are shown in fig. 3.

Algorithm for Film Artifacts Removal

Step 1. The MRI image is read and stored in a two dimensional matrix.
Step 2. Four corners of MR image are tracked. The peak threshold is set for removing white labels in selected areas.
Step 3. Flag value is set to 1.
Step 4. Pixels with intensity value greater than or equal to 200 are selected.
Step 5. The flag value is set to zero if the intensity value is greater than 200.
Step 6. Otherwise skip to the next pixel.

C. Skull Removal

Skull removal from Brain MRI is an imperative signpost. The technique that is to be used to remove skull, must be effective, efficient, reliable and fully automated. Many techniques are available for skull removal. Some techniques are effective but very inefficient and vice versa. In Brain MRI there is a particular intensity of the back-ground that appears before brain image. Unfortunately in brain MRI, the same intensity is appeared as a part of the brain. And this appearance is a false background. So in this scenario that algorithm would be unable to distinguish between the original back ground and the false background. Eventually the area around the false background will also be eroded, which causes distortion in the brain tissues along with the skull. To avoid such distortion, proposed method employs morphological techniques. The morphological techniques are effective and fully automated but inefficient way of skull removal. Erosion and Dilation are two fundamental operations in mathematical morphology. These are applied to binary images, but can also be applied on grayscale images [17]. Morphological image opening i.e. erosion followed by dilation is performed on film artifacts removed MR image. Erosion is a technique which uses background and the foreground for the processing. Erosion detects the small gap between skull and brain, refer fig. 4.

Algorithm for skull removal

Step 1. A film artifact removed image is taken as input.
Step 2. Threshold is set to convert the gray scale image into binary image.
Step 3. Appropriate structuring element is selected with suitable parameters.
Step 4. Morphological opening is performed.

IV. PROPOSED METHODOLOGY- CLUSTERING

Clustering is vital technique used in pattern recognition and region based segmentation approaches. Group of objects which have some similar features between them and are dissimilar to the objects belonging to other groups is defined as cluster. Clustering can be defined as an unsupervised learning process of organizing objects into groups whose members possess some similar feature. Proposed methodology follows Hierarchical and K-means clustering methods to detect tumor.
A. Hierarchical Clustering

A method of cluster analysis which looks for building a hierarchy of clusters is termed as Hierarchical clustering. Two types of Strategies for hierarchical clustering are Agglomerative and Divisive clustering. In agglomerative clustering each data item is considered as a cluster and clusters are recursively merged to yield a good clustering. In divisive clustering the entire data set is regarded as a cluster, and then clusters are recursively split to yield a good clustering. Agglomerative clustering is a “bottom up” strategy while divisive is a “top down” strategy. Proposed work follows divisive i.e. top down strategy, refer fig. 2. This is a semi-automatic method as user has to enter the number of iterations to be performed.

Input MR Image is analyzed to study intensity values of all pixels. To find clusters in image, mean of image is calculated and all pixels are compared with mean value. Some pixels having intensity value less than mean form one cluster while some pixels having intensity value greater than mean form other cluster. Thus image is now divided into 2 clusters. Since for T2 weighted image, tumor region falls in white color shade, cluster of further interest is of pixels having intensity values greater than mean. This cluster is selected and again same procedure is repeated on it. The number of times procedure to be repeated is taken from user. Results are shown in fig. 5.

Algorithm for hierarchical clustering

Step 1. A skull removed image is taken as an input. Consider whole image as one cluster.
Step 2. Number of iterations to be performed is taken from user.
Step 3. The image is divided into two clusters by knowledge of most dissimilar points in the image.
Step 4. Repeat step 3 for each cluster.
Step 5. A tree like structure is formed. Repeat step 3 until it reaches level 4.

B. K-Means Clustering

Unsupervised learning techniques require prior knowledge such as number of clusters, probability distribution etc. K-means clustering is used to divide data into certain number clusters; K. Algorithm assigns each of data points to one of the K clusters. This assignment is based on Euclidean distance between data point and mean of the cluster. Data point is assigned to particular cluster whenever Euclidean distance is minimized.

Proposed methodology uses K=4 clusters for T2 weighted MR images. In T2 weighted MR brain images there are 4 clusters classified on basis of intensity values of pixels. These clusters are nothing but White matter (WM), Gray matter (GM), Cerebrospinal fluid (CSF) and background of MR image. WM, GM, CSF are constituents of brain. By knowledge of mean of each cluster, each pixel in image is assigned nearest cluster using Euclidean distance. This procedure is repeated until we get best clusters. Results are shown in fig. 6.

Algorithm for K-means clustering

Step 1. Initial clusters are chosen according to the gray values of the pixels.
Step 2. Center of each cluster is calculated. This center is used as the new value for the given cluster.
Step 3. Each pixel in the input image is compared against the initial cluster centers using the minimum Euclidean distance.
Step 4. Once the new cluster values have been determined; each pixel is assigned the nearest cluster.
Step 5. The algorithm continues until pixels are no longer changing the cluster which they are associated with
or until none of the cluster values changes by more than a set small amount.

V. RESULTS

![Figure 3. Results of Film artifacts removal (a) Original T1 weighted MR image, (b) Film artifact removed image](image1)

![Figure 4. Results of skull removal for T2 weighted image (a) Film artifacts removed image, (b) Skull removed image](image2)

VI. CONCLUSION AND FUTURE WORK

This research was conducted for accurate detection of brain tumor using clustering techniques. To examine
proposed methodology, MR brain images were taken and processed through clustering techniques thus giving
efficient end results for detection of tumors. These techniques give efficient and effective results as compared
to previous researches. These techniques were applied on various images and results were extraordinary.
Proposed research is easy to execute, scalable and can be modified easily.
Future work is to extend our proposed method for Pseudo colour based segmentation of 3D images,
segmentation based on fuzzy logic, tumor detection using support vector machines (SVM) and using wavelet
technique.
Figure 5. Result of Hierarchical clustering (a) Level 1 clustered (b) Level 2 clustered (c) Level 3 clustered (d) Level 4 clustered

Figure 6. Results of K-means clustering (a) Input image, (b) Gray matter, (c) CSF, (d) Extracted tumor
ACKNOWLEDGEMENT

We would like to thank the “Whole Brain Atlas” and “Mangalam Hospital & Reva Labs” for providing brain MR images.

REFERENCES