Gray Matter and White Matter Segmentation from MRI Brain Images Using Clustering Methods

Deepa V, Benson C. C, Lajish V. L

Department of Computer Science, University of Calicut, Kerala, India-673635

Abstract - Brain is the most complex and master organ of the human body. Quantitative analysis of anatomical brain tissues such as GM and WM is important for the clinical diagnosis and therapy of neurological diseases. Brain imaging is a widely used method by the doctors and clinicians for the representation of human brain. Magnetic Resonance Imaging (MRI) is a recently emerged method for brain imaging. Accurate segmentation of the brain tissues from the MR images is useful in neuro-diagnosis and neurosurgery. It will help the doctors for the diagnosing of some complex diseases such as Epilepsy, Stroke, Alzheimers disease, brain tumor, brain infection and multiple sclerosis. This paper proposes an efficient method for the automatic segmentation gray matter and white matter regions from MRI brain images using Fuzzy c-Means (FCM) and Kohonen means (k-Means) clustering methods. We implemented the clustering of gray matter and white matter using intensity values and statistical feature based values. Finally the results are compared with the manually marked ground truths using some standard accuracy measurement coefficients. The experimental result shows that statistical feature based clustering produces more prominent result than intensity based clustering.

Key Words: Gray Matter, White Matter, k-Means, Fuzzy c Means.

1. INTRODUCTION

The human brain controls and systemizes the overall functions and activities in the human body. Based on embryonic development, we can anatomically divide brain into three main regions: the forebrain, midbrain and hindbrain. The forebrain (prosencephalon) is made up of thalamus, hypothalamus, cerebrum, and pineal gland and other features. The centre of the brain which is called midbrain (mesencephalon), is composed of a portion of the brainstem. The hindbrain (rhombencephalon) consists of the remaining brainstem as well as our cerebellum and pons. Brain cells can be broken into two groups: Neurons, helps in the communication and processing within the brain and Neuroglia, or glial cells support and protect the neurons [1-3]. The structural complexity of the brain made the researchers and medical practitioners to work on to study its function and the diagnosis of diseases. The brain can be sectioned in three planes. Each section provides a different view of the internal anatomy of the brain: sagittal (lateral plane), coronal (frontal/transverse), and horizontal (axial view).

Procedures and techniques used to create images of the human body are called medical imaging. This helps us to visualize internal body structures for precise diagnosis of a wide range of anatomical and physiological disorders. Medical imaging plays a vital role in the field of study of brain. Brain imaging is the widely applicable method for diagnosing many brain abnormalities like brain tumour, stroke, paralysis, Epilepsy, Alzheimers disease etc. Recently there are many types of high resolution imaging techniques are emerged such as MRI scan or CT scan [4]. MRI is to study the finer details in the internal structure of the body. Pathological tissue can be easily discriminated from normal tissue by using Magnetic Resonance Imaging, thus help doctors in identifying brain related disorders. The main advantage of using MRI instead of other imaging techniques is that MRI does not emit any harmful radiation to the human body. MRI makes the use of the property of nuclear magnetic resonance (NMR) to image nuclei of atoms inside the body. An MRI machine uses a magnetic field and radio waves to create detailed images of the brain. Brain tissue can be broken down into two major categories gray matter and white matter, which are the important subject of study in brain imaging. Gray matter consists of mostly unmyelinated neurons, most of which are interneuron. Nerve connections and its processing are the two important areas related to gray matter. Gray matter is composed of neural and glial cells which controls the brain activity. These neuro-glia or glia has gray matter nuclei which is located deep within white matter. White matter is made up of mostly myelinated neurons that connect gray matter to each other and to the rest of the body parts. The information highway of the brain is the white matter which speeds up the connections between distant parts of the brain and body. White matter fibre consists of many elinated axons which connect the cerebral cortex with other brain regions [5].
Determination of GM and WM volume from brain magnetic resonance images (MRI) has become an important measurement tool for identifying different diseases. Calculation of changes in gray and white matter volume in Migraine and Huntingtons patients is very significant for the clinical diagnosis [6-8]. In multiple sclerosis, it is becoming increasingly clear that there is gray matter pathology. Myelin sheath undergoes extensive damage when multiple sclerosis occurs [9]. Similarly gray and white matter changes are observed in patients with Alzheimers disease with extensive gray matter atrophy [10].

In this paper we used clustering methods for extracting gray matter and white matter from MRI brain images. Clustering could be the process of organizing objects into groups whose members are similar in some way. To do this, some important clustering algorithms such as k-Means (Kohonen means) and Fuzzy c-Means methods were implemented in this paper. The k-Means and Fuzzy c-Means are clustering algorithms, which partition a data set into clusters according to some defined distance measure. We applied the clustering algorithms using intensity based values and statistical feature based values. Finally the results are compared with the manual segmentation ground truths using segmentation accuracy checking measures such as Dice and Tanimoto coefficients.

The rest of the paper organized as follows, literature survey explained in the next section, followed by methodology as third part, IVth section includes comparison and evaluation and conclusions are described in the last part.

2. LITERATURE SURVEY

Segmentation is an important stage in the medical image processing. In the literature, we can see many methods for segmenting region of interest from the MRI brain image. Resmi Aet al. proposed a novel method for extracting gray matter and white matter from MRI brain images. They extracted the pathological structures from the images using spatial domain filtering techniques which mainly involved morphological filtering, correlation filtering and local adaptive thresholding [13]. Kaushik K S et al. proposed another method for segmentation of the white matter from the fMRI brain images. They used histogram based double thresholding method for noise removal and then performed white matter segmentation on the basis of a threshold value. The threshold value is determined by using a global thresholding method called Otsu's algorithm. Otsu's method selects the threshold by minimizing the within-class variance of the two groups of pixels separated by the thresholding operator [14].

Andrew Melbourne et al. introduced a novel method for the measurement of white matter structure and features of the cortical surface in the neonatal brain based on incorporating a Maximum a Posteriori Expectation Maximization (MAP-EM) based probabilistic segmentation technique, which includes intensity non-uniformity (INU) correction, spatial dependence via a Markov Random Field (MRF) and implicit correction of PV containing voxels [15]. Mainak Sen et al. proposed an efficient method for Cerebral White Matter using Probabilistic Graph Cut Algorithm [16].

In this paper we used clustering methods for extracting gray matter and white matter from MRI brain images. In the literature we can see many applications of clustering methods in the medical images. k-Means [17] and Fuzzy c-Means [18] are the two important clustering methods used in the medical image processing field.

In the literature we can see several applications of clustering algorithms. Nguyen van Huanet al. proposed an efficient approach for locating eye features in color images based on the unsupervised K-means clustering. This method detects the iris by unsupervised K-means clustering on the feature spaces of compensated red and green color channels [19]. The authors also performed experiments on a collection of eye images extracted from FERET facial database and self-collected images showed a promising performance. Pan Ng and Chi-Man Pun used K-means clustering methods for skin color segmentation. Here author improved the traditional skin classification by combining both color and texture features for skin segmentation [20].

Xuechen Li et al. implemented a method for segmentation of liver from CT scanned images using fuzzy c-mean clustering and level set. The Basic steps involve contrast enhancement of original image followed by spatial fuzzy c-mean clustering combining with anatomical prior knowledge is employed to extract liver region automatically [21].

In this paper we also used some statistical features for the clustering purpose. Feature extraction is the technique of extracting specific features from the pre-processed images of different abnormal categories in such a way that the within - class similarity is maximized and between - class similarity is minimized. Earlier research works report many feature extraction techniques employed for medical image processing. Ebrahim Aghaiai and D. C Gharpure extracted different features of an image and then clustered
the image based on features using Fuzzy c-Means and K-means [22]. They extracted the statistical parameters such as variance, standard deviation and energy from the image. Other feature extraction methods from the medical images are given in the references [23, 24].

3. MATERIALS AND METHODS

Separation of gray matter and white matter from MRI brain image is an important and essential task in the identification of various abnormalities and disorders of brain. This proposed work extract gray matter and white matter regions from MRI brain images using Fuzzy c-Means (FCM) and kohonen-Means (k-Means) clustering methods. The clustering is performed using intensity values and statistical features. Block diagram of the proposed work is represented in Fig-1.

Fig -1: Block Diagram of the proposed work

3.1 Dataset

We have used a total of 50 DICOM MRI brain images of axial plane for experimental purpose. MRI data was provided by Baby Memorial Hospital, Calicut. All the images used for the experiment are of T2 sequences. T2 images were chosen for the work since sequence provides optimum differentiation between different soft tissues of brain. On T2 images, gray matter is light gray and white matter is dark gray and CSF is white. T2 pulse sequences even provide sufficient contrast to segment the images into four compartments (i.e., white matter, gray matter, CSF and abnormal tissues) [28]. Image dimension of 512 X 512 has been used for the proposed work. Clustering of GM and WM from MRI images of the brain was done by using MATLAB toolkit.

3.2 Pre-Processing

Image pre-processing is one of the challenging and crucial factor in computer-aided detection systems. It is the process of converting an acquired image into a useful format. The main objectives of pre-processing are the suppression of redundant and irrelevant data before processing into an application. Pre-processing helps in the correction of noisy, distorted and degraded images. Pre-processing techniques like de-noising and resolution enhancement are used to improve the image quality and features, relevant for further processing and analysis. In this paper we applied three pre-processing steps- image de-noising, image enhancement and Skull stripping.

Medical images are prone to a variety of types of noise which affects the quality of the image. It is the outcome of errors in the image acquisition process so that the resulting pixel values do not reflect the true intensities of the real scene. Examples of such noises are salt and pepper noise, additive white Gaussian noise and speckle noise. Amongst these, speckle noise is an inherent property of medical image. The speckle noise is created due to in-homogeneity of magnetic field, exertion of temperature, motion of the tissue, etc. [26]. Speckle noise reduces the resolution and contrast of the image, thus reducing the diagnostic value of the imaging modality. So speckle noise reduction is an essential pre-processing step, whenever MR imaging is used for medical imaging.

Median filter is commonly used for reducing speckle noise due to their robustness against impulsive type noise and edge preserving characteristics. In the median filter all the pixel value are sorted from the surrounding neighborhood into numerical order and then replaces the pixel under consideration with the middle pixel value. Median filtering is similar to using a mean (averaging) filter, in that each output pixel is an “average” of the neighborhood pixel values. The median filter often does a better job than the mean filter in preserving useful detail in the image. The two main advantages of median filter over the mean filter is that it is much more robust and much better at preserving sharp edges than the mean filter. Review of conventional image enhanceemen filtering technique such as Median filter is presented and its importance in the removal of speckle noise from MR images is also detailed in the literature [27, 28]. Output of de-noising is given in the Fig-2.

One of the main problems associated with the MR images are its low contrast. It is due to noises and other magnetic field in-homogeneities. So for better segmentation result we have to enhance the contrast of the input images. Another problem is the presence of skull. Though skull is an important part in the human body, it is not needed for diagnosing using MR Images of brain. Here we have used a mathematical morphology based image enhancement and skull stripping method proposed by Benson. C. C and Lajish, V. L [29]. Output of image enhancement and skull stripping are given in Fig- 3.
3.3 Clustering

Extraction of gray matter and white matter from MRI brain images plays an important role in the field of medical science. In our work we have applied k-Means and Fuzzy c-Means clustering approaches for performing GM and WM segmentation. The objective of clustering is to make a decision on the intrinsic grouping of unlabeled data. An excellent clustering or a good clustering is independent of the final aim of clustering. The outcome of clustering absolutely depends on user’s needs. Clustering application includes pattern recognition, image analysis, data mining, statistical analysis etc. There are mainly four types of clustering algorithm: Exclusive Clustering, Overlapping Clustering, Hierarchical Clustering, and Probabilistic Clustering.

In k-Means clustering data is grouped in an exclusive way. It follows exclusive clustering algorithm where certain datum belongs to a definite cluster which could not be included in another cluster. Fuzzy c-Means is an overlapping clustering algorithm. It uses fuzzy sets to cluster data, so that each point may belong to two or more clusters with different degrees of membership. There will be appropriate membership value associated for each datum in Fuzzy c-Means. The distance measures between data points play an important role in clustering. If the data instance vector components are all in the same physical units then it is possible that the most common geometric measure, Euclidean distance metric is sufficient to successfully group similar data instances.

Fig -2: Removal of Speckle noise using Median filter

Fig -3: Result of image enhancement and skull stripping

k- Means Algorithm

K-Means is one of the simplest unsupervised clustering algorithms, proposed by MacQueen in 1967 and was originated from the field of signal processing [30]. K-Means follows a numerical, unsupervised, non-deterministic and iterative method for partitioning images into k clusters. Selection of k can be done manually, heuristically and randomly. At first the pixels are clustered based on their color, intensity, texture, and location or a weighted combination of these factors where the clustering process is accomplished. This algorithm is guaranteed to converge but provides only feasible solution. The quality of the solution depends on the initial set of clusters and the value of K. A good clustering will produce high quality clusters with high intra-class similarity and low inter-class similarity. The main problem with the k-means algorithm is the selection of initial centroid. If the initial centroid selection is not proper, it will affect the result of the clustering. So we selected a method proposed by Dost Muhammad Khan and Nawaz Mohamudally [31]. They calculated the difference between the maximum intensity and minimum intensity and the result is divided with the number of clusters, K.
The k-Means algorithm is described below.

1. Let X = \{x_1, x_2, \ldots, x_n\} be the input data set.
2. Choose the number of desired cluster value as K from input MR image.
3. Compute the K initial centroids by using the range method 
   \[
   \left( \frac{x_{\text{max}} - x_{\text{min}}}{K} \right)
   \]
   where \(x_{\text{max}}\) and \(x_{\text{min}}\) represents maximum and minimum image intensity.
4. Compute the cluster centroid \(m_i\) (mean) of each of the \(\sum_{i=1}^{n} x_i\)
   K clusters where \(m_i = \frac{1}{n} \sum x_i\), where \(k=1..K\).
5. Partitioning the input data points into K clusters by assigning each data point \(x_j\) to the closest cluster centroid \(m_i\) using the selected distance measure i.e., Euclidian Distance (ED).
6. ED (i) = argmin \(||x_i - m_k||^2\) where i=1...N.
7. Allocate each pixel to cluster with closest centroid.
8. Re-estimate the centre by calculating mean.
9. Repeat the process from step 3 until the centre doesn’t change.
10. Separate Image into K sub images according to clustered indexed Image.

**Fuzzy c-Means Algorithm**

Fuzzy c-Means algorithm is an iterative clustering method that produces an optimal number of partitions, c by minimizing the weighted within group sum of squared error objective function (WGSS). Fuzzy c-Means clustering algorithm was first reported in the literature by Joe Dunn in 1973[32]. Later it was improved by Jim Bezdek, Cornell University in 1981[33]. Equation to find WGSS is given below.

\[
Q(U, V) = \sum_{i=1}^{n} \sum_{j=1}^{c} M_{ij}^c ||x_i - r_j||^2
\]

Where \(||x_i - r_j||\) is the Euclidean distance between \(i^{\text{th}}\) data and \(j^{\text{th}}\) cluster center, \(m\) is the Fuzziness index which is any real number greater than 1.

Let X = \{x_1, x_2, x_3, \ldots, x_n\} be the set of n dimensional data points and R = \{r_1, r_2, r_3, \ldots, r_c\} be the set of centers. Main objective of fuzzy c-means algorithm is to minimize the objective function.

The algorithm is given below.

1. Initialize Degree of membership matrix \(M = [M_{ij}]\) matrix, \(M(0)\)
2. At k-step: calculate the cluster centre vectors \(R(k) = [R_i] \) with \(M(k)\)
3. Compute the fuzzy centres \(R_i\) using equation given below:
   \[
   \frac{\sum_{j=1}^{c} M_{ij}^{c}}{\sum_{j=1}^{c} M_{ij}^{c}}
   \]
4. \(R_i=\forall i=1, 2, \ldots, c\) is the cluster centers
5. Update \(U(k)\), \(U(k+1)\) where \(U = (M_{ij})_{i=1\ldots c}\) is the fuzzy membership matrix.
6. Degree of membership of \(x_i\) to the cluster \(j\) is given as,
   \[
   \frac{1}{\sum_{j=1}^{c} \frac{||x_i - r_j||^2}{||x_i - r_k||^2}}
   \]
7. The process stops if \(\max_{ij} ||M(k+1) - M(k)|| < \partial \) or minimum Q is obtained, where \(\partial\) is the termination criterion between [0, 1]
8. Otherwise go to step 3.

**4. RESULTS AND EVALUATION**

Extraction of gray matter and white matter from MRI brain images is a challenging task in the medical science field. Here we have used k-Means algorithm and Fuzzy c-Means algorithm for the segmentation purpose. We separately applied clustering methods on the basis of intensity values and statistical feature based values. In statistical feature extraction, the following features are investigated: mean, variance, standard deviation, kurtosis, skewness and range. Among them clustering using mean produces good result. So we have only considered the mean feature and other statistical features are excluded.

The ground truth of the images are created using ITK-Snap tool [34]. It is open source software that contains innovative tools for manual outlining and quality control. The performance of the algorithms is evaluated using the Dice [35] and Tanimoto [36] coefficients. These are two standard measures used to compare the segmentation accuracy.

\[
\text{Dice Coefficient} = \frac{2|A(S) \cap A(G)|}{|A(S)| + |A(G)|}
\]

\[
\text{Tanimoto Coefficient} = \frac{|A(S) \cap A(G)|}{|A(S)| + |A(G)| - |A(S) \cap A(G)|}
\]

Where A(S) represents area of the gray matter and white matter that we have calculated using proposed algorithm. A(G) represents area of the gray matter and white matter that we marked as ground truth. Figure 4 to Figure 7 represents sample results of the algorithms.

The performance of the algorithms is given in the tables. Table.1 represents the accuracy of the k-Means algorithm using Dice coefficient and Tanimoto coefficients and Table.2 represents the accuracy of the Fuzzy c-Means
algorithm using Dice and Tanimoto coefficients. Graphical representation of the results is given in the Figure 8. From the tables we can see that statistical feature based clustering produces more good results than the intensity based clustering.

Table-1: k-Means performance comparison

<table>
<thead>
<tr>
<th>Method</th>
<th>k-Means (Average performance)</th>
<th>GM</th>
<th>WM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Dice</td>
<td>Tanimoto</td>
</tr>
<tr>
<td>Intensity based</td>
<td>0.709658</td>
<td>0.723641</td>
<td>0.713987</td>
</tr>
<tr>
<td>Statistical based</td>
<td>0.803633</td>
<td>0.802788</td>
<td>0.808185</td>
</tr>
</tbody>
</table>

Table-2: Fuzzy c-Means performance comparison

<table>
<thead>
<tr>
<th>Method</th>
<th>Fuzzy c-Means (Average performance)</th>
<th>GM</th>
<th>WM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Dice</td>
<td>Tanimoto</td>
</tr>
<tr>
<td>Intensity based</td>
<td>0.714960</td>
<td>0.696072</td>
<td>0.791050</td>
</tr>
<tr>
<td>Statistical based</td>
<td>0.852144</td>
<td>0.821170</td>
<td>0.864319</td>
</tr>
</tbody>
</table>

Fig-4: Results using intensity based k-Means clustering

Fig-5: Results using statistical based k-Means clustering

Fig-6: Results using intensity based FCM clustering
5. CONCLUSIONS

The gray and white matter segmentation of brain image is vital in identifying disorders and treatment planning in the field of medicine. In this work we extracted gray matter and white matter using two well-known clustering algorithms: k-Means and Fuzzy c-Means algorithms. In both cases we performed intensity based and statistical feature based clustering. In statistical features we have used mean, variance, standard deviation, kurtosis, skewness and range. Among them only mean feature produces the prominent results. Experimental results shows that the segmentation using statistical based (mean) clustering gives improved performance metric and improved classification accuracy rather than intensity based segmentation. Our work investigated features based on first order statistical parameters that give very less number of distinguishable features for classification of MR images. In the future, we can use second order statistical features for the better segmentation of GM and WM in MR images.

REFERENCES

[8] Andrea Ciarmiello, Milena Cannella, Secondo Lastoria, Maria Simonelli, Luigi Frati, David C Rubinszttein and Ferdinando Squitieri, “Brain white-


Journal of Information Processing and Management, Volume 1, Number 1, July 2010.


BIOGRAPHIES

Deepa V has completed her master of computer applications from Avinashilingam deemed University in the year 2000. She is an MPhil research scholar working in the area of medical imaging under the guidance of Dr. Lajish V.L.

Benson, C. C earned his Masters in Computer Applications from the University of Calicut in 2011 and specialized in Medical Image Processing. He is an active research scholar working in the area of MR Image analysis under the supervision of Dr. Lajish V.L.

Lajish V.L. has been associated with University of Calicut, Kerala, as Head of the Department of Computer Science. He has worked as Scientist R&D in TCS Innovation Labs, Tata Consultancy Services Ltd. Mumbai, prior to joining the University. His prime areas of research include Digital Speech and Image Processing, Pattern Recognition algorithms, Indian language script technology solutions for masses. After his Masters in Computer Applications from Vellore Institute of Technology, he earned his Ph.D in Computer Science from University of Calicut, Kerala in 2007. He is a senior life member of International Association of Computer Science and Information Technology.