Indian Journal of Science and Technology, *Vol 9(46), DOI: 10.17485/ijst/2016/v9i46/106495, December 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Comparison and Evaluation of Segmentation Techniques for Brain MRI using Gold Standard

M. Akhil[1*], R. Aishwarya[1], Vandhana Lal[1] and Shanthi Mahesh[2]

[1]Atria Institute of Technology, Anand Nagar, Bangalore - 560024, Karnataka, India
[2]Department of ISE, Atria Institute of Technology, Anand Nagar, Bangalore - 560024, Karnataka, India

Abstract

Objective: Automated segmentation is an active research for medical images. Accuracy of automated segmentation methods plays a vital role during brain image analysis. Segmentation being an important area of research, determining its performance is also important. Gold Standard is required for comparison during segmentation evaluation. **Method**: The Gold Standard for segmentation of medical images is the manual drawing of region of interest. This manual tracing is performed by experts (radiologists). The deviation of segmentation when compared with the experts and the quality of segmentation are inversely proportional. **Analysis**: The quantitative methods indicate the performance of the segmentation methods when compared with Gold Standard. Evaluation metrics mostly fall into three categories: Area Based Evaluation method (Dice coefficient, Jaccard Coefficient, Relative Volume Difference, Volume Overlap error), Surface Evaluation type (Average Symmetric Surface Distance, Root Mean Square Symmetric Surface Distance, Scatter Plot) and Specificity, Sensitivity and Accuracy.

Keywords: Gold Standard, Segmentation, MRI, Manual Segmentation, Automated Segmentation, Evaluation Metrics

1. Introduction

A rapid advance in Magnetic Resonance Imaging (MRI) has enabled non-invasive exploration of the human brain with a high level of detail. The implementation of image processing for large longitudinal brain imaging poses difficulty in the development, implementation and validation of different components. Methods that are employed for cross-section studies cannot detect high sensitive sublet longitudinal changes. Recent studies based on volumetric analysis of the brain involves cross-sectional data and attempt to find the measurement of either a part or total volume from the area of interest[1–5]. Imaging studies that employ quantitative volumetric methods obtain more accurate results than qualitative methods. Quantitative neuroanatomical methods vary from manual, local operations in 2D images to automated etc[6]. An extensive literature exists on stereological methods[7,8] that are appropriate only for volumetric measurements of large partitions and detection of small changes by fine sampling requires excessively high amount of work.

The brain is a soft, delicate, non-replaceable and spongy mass of tissue. Brain tumour is a mass of aberrant cells that grow in the brain. Tumours can harm the healthy cells directly or indirectly by pervading the other parts of the brain by causing swelling, inflammation and pressure within the skull. The main aim of this work is to compare and evaluate the tumours using brain MRI. MRI is a medical imaging technique used to represent the anatomy and physical processes of body. MRI scanners use strong radio waves, magnetic fields and gradients to images of the body. MRI displays a digital depiction of tissue characteristic that can be obtained in any tissue plane. MRI images are best described as slices through the brain. The added advantage of MRI is that the slice is through both horizontal and vertical planes. The complete visualization image of the human body aids the surgeons to perform operations without the need for open surgery. More specifically, ease of treatment increases as the surgeon is able

**Author for correspondence*

to see the tumour and segment it. Because of the variety in the tumours shapes, location and image intensities segmentation is a difficult and challenging task.

2. Materials and Methods

Image Analysis comprises of steps like re-slicing of the data, purging of extra-cranial tissues, segmentation of different types of tissues, normalization to a standard coordinate space and regional quantitative analysis. Figure 1 illustrates the overall procedure. In this section, each component is described in a greater detail.

The concept behind this work is based on four main points:

- A brain MRI is considered.
- The region of interest (tumor) is manually segmented from the original image.
- The region of interest (tumor) is segmented using different automated segmentation techniques.
- The manual and automated segmented images are compared and evaluated using the evaluation metrics.

2.1 Manual Segmentation

Manual segmentation refers to the process whereby an expert transcriber segments and labels a speech file by hand, referring only to the spectrogram and/or waveform. For manual segmentation method, the experts of brain tumor must master the information presented in the brain tumor images and some additional knowledge such as anatomy because manual brain tumor segmentation aims to manually draw the boundaries of the brain tumor and paint the regions of anatomic structures with different labels.

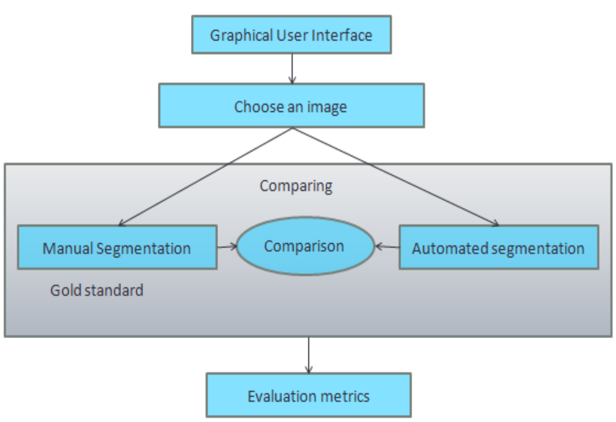


Figure 1. Architectural diagram.

Manual segmentation is a reliable but time consuming task. It is highly based on large inter and intra observer variability that produces a degradation in credibility of the analysis. Therefore, automated segmentation methods are used to conduct more reliable, accurate and robust analysis[10].

2.2 Automated Segmentation

For automatic brain tumor segmentation method, the computer determines the segmentation of brain tumor without any human interaction. In general, an automatic segmentation algorithm combines artificial intelligence and prior knowledge. With the development of machine learning algorithms that can simulate the intelligence of humans to learn effectively, the study of automatic brain tumor segmentation has become a popular research issue.

2.2.1 Thresholding

A straightforward method of segmentation is the thresholding method. This method works on the principle of turning gray-scale image into a binary image using a threshold value. The key is to select the appropriate threshold value. Some of the popular methods used are Otsu's method, maximum entropy and k-means clustering[9].

2.2.2 Histogram

Histogram methods are very effective when compared because they need only a single pass through the pixels. Using the pixels from the image, a histogram is figured. The peaks and valleys in the figure are used to locate clusters with the help of intensity as a measure.

2.2.3 Confusion Matrix

It is a specific table layout that depicts the performance of a classifier on a test data for which true values are known. The matrix is simple to understand.

True Positive (TP) - Measures the proportion of positives that are correctly identified.

True Negative (TN) - Measures the proportion of negatives that are correctly identified.

False Positive (FP) - Result that incorrectly indicates about a particular condition being present.

False Negative (FN) – Result that incorrectly indicates about a particular condition being absent.

2.2.4 Evaluation Metrics

Following are the derivations of the metrics using the four basic cardinalities from the confusion matrix, namely TP, FP, TN and FN.

Dice coefficient: It is the measure of extent of spatial overlap between two binary images. It is mostly used to report the performance of segmentation. It gives more weightage to instances where the images agree.

$$DICE = \frac{2\left|S_g^1 \cap S_t^1\right|}{\left|S_g^1\right| + \left|S_t^1\right|} = \frac{2TP}{2TP + FP + FN}$$

Jaccard Index: It is the metric evaluation that allows to give the similarities of the segmentation comparing the Ground-Truth and the segmented image.

$$JAC = \frac{\left|S_g^1 \cap S_t^1\right|}{\left|S_g^1 \cup S_t^1\right|} = \frac{TP}{TP + FP + FN}$$

Sensitivity: It is also known as True Positive Rate (TPR), calculates the proportion of positives that are identified correctly.

$$Sensitivity = TPR = \frac{TP}{TP + FN}$$

Specificity: It is also known as True Negative Rate (TNR), calculates the proportion of negatives that are identified correctly.

$$Specificity = TNR = \frac{TN}{TN + FP}$$

Precision: A measure of degree to which the same result would be produced over different segmentation sections.

$$Precision = PPV = \frac{TP}{TP + FP}$$

F-measure: It is the measure of a test's accuracy. It considers the precision and recall to compute the harmonic mean.

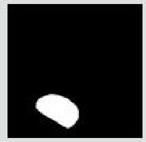
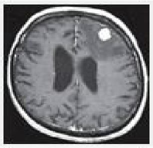
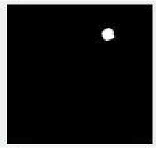
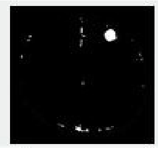
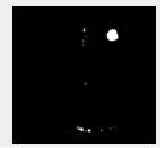
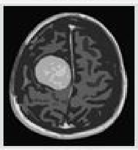
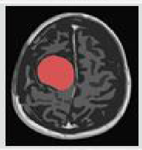
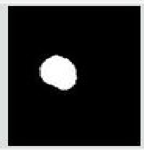
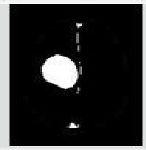
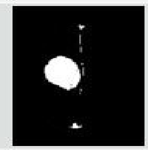
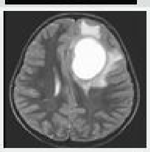
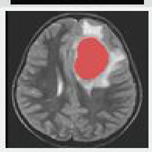
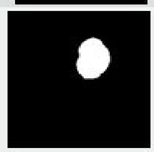
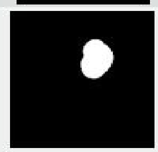
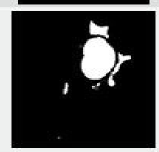
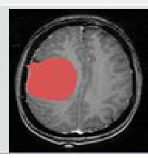
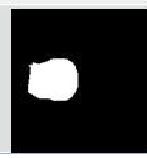
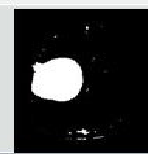
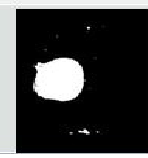
$$FMS = \frac{2 \cdot PPV \cdot TPR}{PPV + TPR}$$

3. Results and Discussion

Table 1. Evaluation of metric values

SAMPLE	METHOD	DICE	JACCARD	COHENS KAPPA	SENSITIVITY	SPECIFICITY	ACCURACY	PRECISION	F-MEASURE
IMAGE 1	GLOBAL	0.9889	0.978	0.9998	0.9845	0.8812	94.72	99.33	98.89
	HISTOGRAM	0.9887	0.9777	0.9998	0.9855	0.8579	94.72	99.2	98.87
IMAGE 2	GLOBAL	0.998	0.996	0.9988	0.9963	0.9583	99.43	99.98	99.8
	HISTOGRAM	0.9989	0.9979	0.9978	0.9986	0.8712	99.43	99.93	99.89
IMAGE 3	GLOBAL	0.9958	0.9917	0.997	0.9954	0.9312	94.85	99.63	99.58
	HISTOGRAM	0.9959	0.9918	0.997	0.9956	0.9302	94.85	99.62	99.59
IMAGE 4	GLOBAL	0.9975	0.9951	0.9991	1	0.9126	94.63	99.51	99.75
	HISTOGRAM	0.9875	0.9754	0.9998	0.9755	0.9984	94.63	99.99	98.75
IMAGE 5	GLOBAL	0.9961	0.9922	0.9992	0.993	0.9916	91.23	99.92	99.61
	HISTOGRAM	0.9968	0.9936	0.9991	0.9954	0.9811	91.23	99.82	99.68

Table 2. Table of segmented images

Original image	Manual segment	Binary	Global	Histogram
				
				
				
				
				

4. Conclusion

In this research work, a comparative study has been successfully performed and verified by a radiologist. All the objectives of the project have been fulfilled, in generating the system. This system has a rich user interface which helps the user to use the system with limited knowledge. MATLAB and image segmentation apps and tools have been the most important from the beginning to the end of the project. As in image processing procedures, the most challenging part was to segment the region of interest. The proposed system's evaluation metrics will benefit the physician and radiologist for their further diagnosis, the treatment procedure and state of tumor monitoring. The segmentation algorithms were tested with the sample images. Global Thresholding and Histogram Thresholding segmentation have achieved Dice coefficient of 99%, Jaccard of 98%, Accuracy of 95% and Precision of 99%. This system has proved successful in obtaining the Accuracy by comparing the manual and automated segmented images.

5. Acknowledgement

The contentment and euphoria that comes along with the successful completion of any task is complete only with the mention of the guidance and encouragement of the people who made this a success.

We would like to express our deep sense of gratitude to our guide Dr. Shanthi Mahesh, Associate Professor, for her valuable guidance, constructive criticism, learned counsel and help throughout the course of the project. We are grateful for her unfailing encouragement and suggestions given to us in the course of the project work would not have been a success.

6. References

1. Andreasen NC, Rajarethinam R, Cizadlo T, et al. Automatic atlas-based volume estimation of human brain regions from MR images. J Comput Assist Tomogr. 1996; 20(1):98–106.
2. Jack JCR, Twomey CK, Zinsmeister A, et al. Anterior temporal lobes and hippocampal Formations: Normative volumetric measurements from MR images in young adults. Radiology. 1989; 172:549–54.

3. Kertesz A, Polk M, Black SE, Howell J. Sex, handedness, and the morphometry of cerebral assymmetries on magnetic resonance imaging. Brain Res. 1990; 530:40–8.
4. Shenton ME, Kikinis R, Jolesz FA, et al. Abnormalities of the left temporal lobe and thought disorder in schizophrenia: A quantitative magnetic resonance imaging study. N Engl J Med. 1992; 327:604–12.
5. Andreasen NC, Ehrhardt JC, Swayze VW, et al. Magnetic resonance of the brain in schizophrenia: The pathophysiological significance of structural abnormalities. Arch Gen Psychiatry. 1990; 47:35–44.
6. Reiss AL, Faruque F, Naidu S, et al. Neuroanatomy of Rett syndrome: A volumetric imaging study. Ann Neurol. 1993; 34:227–34.
7. Roberts N, Cruz-Orive LM, Reid NMK, Brodie DA, Edwards RHT. Unbiased estimation of human body composition by the Cavalieri method using magnetic resonance imaging. J Microscopy 1993; 171:239–53.
8. Pakkenberg B, Boesen J, Albeck M, Gjerris F. Unbiased and efficient estimation of total ventricular volume of the brain obtained from CT scans by a stereological method. Neuroradiology. 1989; 31:413–7.
9. Mustaqeen A, Javed A, Fatima T. An efficient brain tumor detection algorithm using watershed and thresholding based segmentation. IJ Image, Graphics and Signal Processing. 2012; 10:34–9.
10. Available from: www.phonetik.unimuenchen.de/forschung/BITS/TP1/Cookbook/node134.html