International Journal of Advance Research In Science And Engineering http://www.ijarse.com
IJARSE, Vol. No.4, Issue No.02, February 2015 ISSN-2319-8354(E)

ACCURACY AND CONSISTENCY OF AUTOMATED BRAIN TUMOR SEGMENTATION IN GRAY MATTER MAGNETIC RESONANCE IMAGES

S.Ramakrishnan[1], K.KharmegaSundararaj[2] , K.Vinoth Raja[3]

PG Student [1], Associate Professor [2],Assistant Professor[3]

Department of Computer Science and Engineering,

[1,2]PSN College of Engineering and Technology, Tirunelveli.(India)

[3]St. Mother Theresa Engineering college, Thoothukudi. (India)

ABSTRACT

A systemized segmentation of gray matter magnetic resonance imaging data is a essential role for morphometric analysis of the human brain. The proposed approaches re formulate the differences of small area of the brain depend on the method used for both accuracy and consistency of various automated segmentation mechanisms have rarely have been compared. Here the ability of the segmentation mechanisms provided by FCM, FSL and Free Surfer was quantified on simulated and real magnetic resonance imaging data. Here the proposed mechanism FCM, FSL and Free Surfer was boosting and speed up its performance under noisy and unexpected data acquisition terms. Due to the multiple vast scanning the morphometric analyses is not warranted. The proposed approach increase the accuracy, robustness and consistency against the existing analyses mechanism compared with FCM technique reported in the literature.

Keyword: Segmentation, Morphometric, Consistency, Accuracy.

I INTRODUCTION

Automated brain segmentation algorithms segment a structural magnetic resonance imaging (MRI) image into different tissue classes. A MRI image is segmented into gray matter, white matter, and cerebrospinal fluid. Based on this segmentation, methods are available to calculate several neuroanatomical measures, for example gray matter volume, gray matter density, cortical thickness, or cortical curvature. Researchers use these procedures to investigate differences in brain structure between groups or to investigate changes in brain structure over time. Phenomena that are investigated include learning process aphasia, alexithymia, post-traumatic stress disorder, depression, autism, and schizophrenia. The use of systemized segmentation algorithms is desirable, as these algorithms are (i) much faster than manual segmentations and (ii) user independent, that is, they do not depend on expert knowledge in neuroanatomy. However, significant challenges exist as differences in brain structure between groups, or changes within subjects are often very subtle. Therefore the automated segmentation algorithms able to specifically conclude the accurate amount of accuracy and they give related results compare with different images of the same person. At the same time to know the accuracy and reliability of automated segmented algorithms.

Extracting the brain cortex from magnetic resonance imaging (MRI) head scans is one of the important pre-processing steps in analyzing intracranial volumes. Any subsequent analysis, such as tissue segmentation or brain volume and atrophy measurement, will be highly dependent on the robustness and precision of the brain masks generated in the brain extraction step. Brain tumor segmentation process consists of dividing the various tumors, such as solid tumor, edema, and necrosis from the normal brain tissues, such as white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). Although semi automated segmentation by qualified professionals remains advanced in quality to automatic methods, it has two types of drawbacks. The first one is that producing manual or semi-automatic segmentations is very time-consuming, with higher accuracies on more finely detailed volumes demanding increased time from medical experts.

Another problem with manual and semiautomatic segmentations is that the segmentation is subject to variations both between observers and within the same observer. Extraction of brain tumor region requires the segmentation of brain MR images into two segments. One segment contains the normal brain cells consisting of GM, WM and CSF and the second segment contains the tumorous cells of the brain. Correct segmentation of MR images is very important because most of the time MR images are not highly contrast thereby these segments can be easily overlapped with each other. There is a well-built algorithms need to have some Well-organized computer based structure that accurately classifies the boundaries of brain tissues along with minimizing the chances of user communication with the system. Additionally, manual segmentation procedures need at least three hours to complete the tested performance of three commonly used segmentation algorithms, provided by software packages SPM5, FSL, and FreeSurfer.

Within-segmenter analyses revealed volume differences greater than 15%. Between-segmenter comparisons showed an average discrepancy of 24% for real MRI images. Systematic evaluation of different segmentation algorithms. They used simulated brain data that were generated based on varying brain anatomy and varying image quality, as well as real images from nine different individuals and test-retest images of 48 individuals. the manual segmentation has no assurance in tracking the tumor volume during the patient follow-up process and the automatic methods that could achieve a sufficient level of accuracy would be highly attractive for their ability to perform high-throughput segmentation. To provide information to the community regarding which gray matter segmentation procedure they can build upon, we present a systematic evaluation of accuracy and reliability of standard gray matter segmentation algorithms by providing a comprehensive investigation of both segmentation pipelines and within and between-segmented accuracy and reliability using the latest versions of commonly used segmentation algorithms. Importantly, we provide measures of accuracy obtained from real T1 MRI images. To our knowledge this has not been done before in a systematically manner. In our current study we evaluate the segmentation algorithms provide by i)FCM, ii) FSL and iii) Free surfer separately and in combination with mechanisms for skull-stripping and intensity modification. Here we evaluated consistency in terms of coefficient of variation, standard deviation and reliability coefficient of gray matter segmentations on real T1 images.

We determined truthfulness in terms of the Dice coefficient computed for the comparison of ground truth images and corresponding gray matter segmentations in simulated and real T1 brain images. In comparison to earlier studies, our focus was on the real-time investigation of accuracy and consistency in combination with a efficient evaluation of the influence of each processing step on segmentation quality. Now we are able to examine how,

i. The interaction between the pre-processing and segmentation algorithms.
ii. The difference between results from simulated and real T1 images.
iii. Which processing method has the largest influence on segmentation accuracy both in simulated and real T1 MRI images.

II METHODOLOGY

2.1 Data Sets

To examine the accuracy of different segmentation pathways we used (i) 18 real T1-weighted MRI images with expert segmentations of 43 individual structures from the Internet Brain Segmentation Repository (‘‘IBSR data set’’),(ii) twenty simulated T1-weighted MRI images and corresponding discrete anatomical models provided by the Simulated Brain. The IBSR data set consists of high-resolution, T1-weighted volumetric images. These images have been processed by the Center for Morphometric Analysis biasfield correction routines. Experts segmentations of the principle brain structures for white and gray matter images. To determine the reliability of different segmentation pathways we acquired ten MRI images of one individual data’s and iii) the Open Access Series of Imaging Studies (‘‘OASIS data set’’).

Brain Web Data Set			
Data set	**White matter**	**Gray matter**	**CNR**
Image 8	59.18	43.48	14.65
Image 10	58.08	47.74	12.34
Image 12	56.8	46.72	11.09
Image 18	50.97	40.8	10.17
Image 20	5475	43.67	11.08
Image 32	56.09	45.97	10.12

Table 1. Image Quality Parameters of the Brain Web Data Set

We compare these results with those of the standard FCM and several well-know non-fuzzy MRI segmentation techniques found in the literature. We also apply the proposed approach to pathological T1-weighted MRI databases obtained from IBSR and from a local MRI scan center to detect hyper-intense tumors. FreeSurfer does not provide a gray matter map right away, we created gray matter masks from the results of the subcortical segmentations and the cortical parcellations and combined these two masks to get the desired gray matter map.

	SNR		
Data set	**White matter**	**Gray matter**	**CNR**
Image1	47.27	36.28	10.99
Image2	115.98	94.12	22.07
Image3	113.07	80.23	69.17
Image4	69.17	34.42	34.75
Image5	83.92	48.52	35.4
Image6	106.32	58.95	47.37
Image7	39.13	27.52	27.52
Image8	103.84	70.01	33.83
Image9	111.23	71.7	39.64
Image10	40.41	26.67	13.74

Table 2 - Image quality parameters of the IBSR data set

	SNR		
Data set	**White matter**	**Gray matter**	**CNR**
Image1	132.46	86.04	46.68
Image2	137.72	85.81	51.82
Image3	127.06	75.71	49.35
Image4	141.6	89.35	52.4
Image5	125.73	78.81	46.94
Image6	143.17	83.24	83.24
Image7	143.8	84.18	59.72
Image8	144.18	83.67	60.19
Image9	131.53	81.86	49.56
Image10	142.81	64.17	78.67

Table 3- Image quality parameters of the Single Subject data set.

III ALGORITHMS

We used two preprocessing steps in our analyses: skull-stripping and intensity correction. For intensity correction we used the nonparametric nonuniform intensity normalization (N3) algorithm and for skull-stripping (i) the ‘‘watershed’’ (WS) algorithm of FreeSurfer and the BET algorithm of FSL. For gray matter segmentation we used (1) ‘‘Segment’’, (2) ‘‘New Segment’’ FCM, (3) ‘‘FAST’’ and (4) ‘‘FreeSurfer’’

3.1 Skull Removal

The skull and non-brain intracranial tissues like fat, muscle, skin etc., that surround the surface of brain cortex and cerebellum in the brain should be removed. This is necessary to avoid the misclassifications of surrounding tissues, skin and scalp as WM or GM. By removing these objects, non-brain tissues will be removed and only soft tissues will be left. This is done by grey level erosion using a disc-shaped structuring element, which results in the removal of thin connections between brain and non-brain portions. Thus a brain mask with the skull removed was obtained and this method was automated for every image slice. The evaluation of brain tissue classification is a complex issue in medical image processing. A justification technique can be thought of as a combination of two components. One component is a measure of establishing the deviation from the GT, i.e., an evaluation with actual segmentation is needed to assess how actual segmentation deviates from the real one. The second component is the notion of ground truth (GT) against which the result of an algorithm is to be judged.

In the case of segmentation of brain tissue from MRIs, there is indeed a true boundary of the brain tissue for each patient, but it is not known what it is. Approximations to the true boundary can be obtained in the form of the manual segmentation algorithm by experts of neuroanatomy; however, manual segmentation is subject to inter-observer variability and human error. To minimise the influence of these factors while maintaining a means of measuring the segmentation accuracy of the individual ratters, the standard was defined on the basis of independent human observers. Randomly selected images from 20 patients’ image datasets were manually segmented by the radiology expert in the group. Inter-observer variability was also considered. These images were selected as ground truth (GT) images for validation. There are three different measures for evaluating GT with the automatic segmentation algorithm, i.e.,

1. Accuracy evaluation using quantitative techniques;
2. Qualitative techniques.

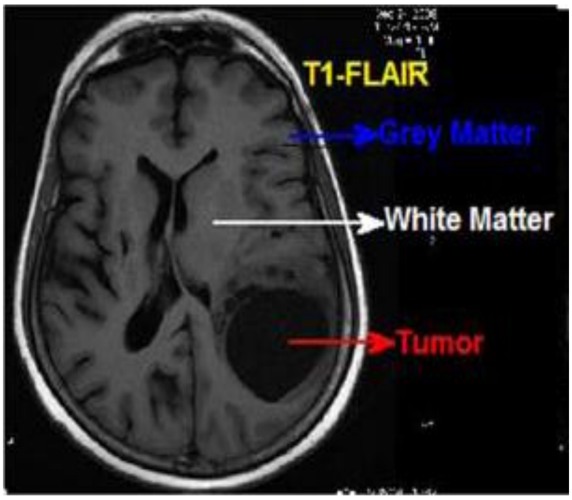

Fig.1 Sample raw data from a patient volume used segmentation.T1-FLAIR

3.2 Accuracy Evaluation

With the GT images in the dataset,to quantify the accuracy of gray matter segmentations, we used the Dice coefficient (DC)three indices were calculated using following things,

$$TI[\%] = \frac{TP}{TP + FP + FN} . 100; \qquad (1)$$

$$P + [\%] = \frac{TP}{TP + 0.5FP}\ 100; \qquad (2)$$

$$PM[\%] = \frac{TP}{GT}\ 100; \qquad (3)$$

Where TP = true positives, i.e., pixels labeled in the GT and by the algorithm; FP = false positives, i.e., pixels labeled as ROI by the algorithm, but not in GT; and FN = false negative. The DC is commonly used to determine accuracy of segmentation methods in neuroimaging settings and is defined as the size of the union of the segmentation result and the ground truth: DC= 2TP/((FP + TP) + (TP + FN)), that is, the set of True Positives (TP) is divided by the average size of the segmentation result (False Positives (FP) + True Positives (TP)) and the ground truth (True Positives (TP) + False Negatives (FN)). A DC of 0 indicates no overlap; a value of 1 indicates perfect agreement. Using the DC, we evaluated the accuracy of the standard implementations of Segment, New Segment, FCM, FASTand FreeSurfer. With regard to the BrainWeb data set, we resliced the gray matter maps produced by the segmentation pathways to the corresponding ground truth images with a trilinear interpolation. Next, we compared the resliced gray matter maps (binarization threshold: p.0.5) and the corresponding ground truth images voxel-wise to calculate the DC. With respect to the IBSR data set, segmentation results could be directly compared to the corresponding ground truth images, because original T1 images and ground truth images had the same resolution. Only in case of FreeSurfer, gray matter maps were again resliced to fit the resolution of the corresponding ground truth images. the percentage ratio between the number of pixels labeled as tumour by GT and the algorithm, and the number of pixels classified as tumour by the algorithm and/or by GT. The value of 100% signifies that there are no FP and FN. Percentage match (PM) index shows the correspondence between the GT and the segmentation algorithm. As a perfomance evaluation, computation time and accuracy of the automatic method with manual segmentation method is compared. The performance of the proposed algorithm on both BrainWeb and IBSR datasets is compared with some reported fuzzy approaches: the standard FCM algorithm and the FCM algorithm with incorporated neighborhood information (NFCM). The latter algorithm is selected because it is one of the most notable FCM-based algorithms imposing spatial constraints.. In addition, the proposed algorithm is compared with the recent non-local FCM family of algorithms We also evaluated the gray matter maps using p.0.10 and p.0.90. To assess the reliability of the five standard segmentation algorithms, we initially used the Single Subject data set. We calculated the variability in segmented gray matter volumes in terms of the standard deviation in mm3 and in terms of the coefficient of variation Next, we calculated the reliability coefficient r for the segmented gray matter volumes measured for the OASIS data set. For this data set, we also computed the average deviation in volume between the first and second scan. Finally, to determine which processing factor had the largest impact on segmentation accuracy, we computed separate univariate, three-way analyses of variance (ANOVAs) with according pairwise comparisons for the BrainWeb data set and the IBSR data set. In these analyses, DC was the dependent variable and Intensity Correction, Skull-Stripping, and Segmentation were therespective factors for repeated measures.

IV RESULTS AND DISCUSSION

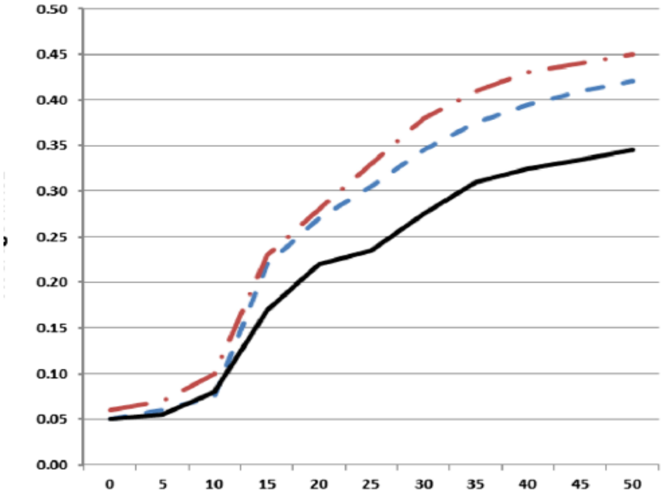

Fig. 2 Performance Analysis of the T1 images.

The BrainWeb data set FAST,FCM, Segment, and New Segment reached an average DC greater than 0.96. FCM achieved the highest accuracy (M=0.9723, SD = 0.0064), while FCM was high consistency (M=0.9474,SD = 0.0067). FreeSurfer showed the lowest accuracy (M=0.8679, SD = 0.0087). On the IBSR data set, as illustrated inpanel B, New Segment reached the highest average DC (M=0.8326,SD = 0.0129). Compared to New Segment, FAST (M=0.7962, SD = 0.0570), Segment (M= 0.8167, SD = 0.0359), and FCM(M= 0.8026, SD = 0.0270) showed slightly lower accuracy. Free- Surfer again showed the lowest accuracy of all segmentation algorithms (M=0.5838, SD = 0.0570).

Thus, the maximum discrepancy between the different segmentation algorithms was 11%. With the exception of FAST, all segmentation algorithms showed very high test-retest reliability on the OASIS data set (all rs .0.97;panel F). FAST, however, only demonstrated a reliability coefficient of 0.90. The corresponding average volume differences between first and second scan were: Segment: 1.2%, SD= 1.1; New Segment: 0.6%, SD = 0.6; FCM:2.0%, SD= 2.2; FAST: 3.3%, SD = 2.6; FreeSurfer: 1.0%, SD= 0.7. in that all segmentation algorithms underestimate the actual gray matter volume. This suggests that, in terms of accuracy, the latest algorithmic advancements have not improved segmentation accuracy significantly. The focus of further investigations should be to determine which brain segmentation algorithm is most accurate for which region of the brain, most importantly, which segmentation algorithm is best suited for the segmentation of corticalareas, and which algorithm provides the most accurate results for sub cortical areas.

V CONCLUSIONS

The main goal of the development of this automated segmentation method is to make segmentation of MR images more practical by replacing manual outlining, which reduces operator time without measurable effect, and to improve reproducibility.Our conclusion address crucial factors that manipulate the quality of gray matter segmentation. Moreover, our results provide guidance in designing state-of-the-art segmentation pathways optimized for individual software settings. Our study emphasizes that comparisons of the results of morphological studies using different segmentation algorithms should be made with great caution. In conclusion, our results suggest that researchers must be aware of the fact that the choice of the segmentation pathway used in a morphometric investigation can easily introduce a ‘‘segmentereffect’’ on the order of 2–4% variability in segmented gray matter volume. Researchers therefore need to optimize their scanning and processing procedure with respect to their individual settings. Before performing a study, the accuracy and consistency of a specific segmentation pathway has to be adequately determined to enable correct interpretation of the results.

REFERENCES

1. Michotte A, Neyns B, Chaskis C, Sadones J and In 't Veld P. Neuropathological and molecular aspects of low-grade and high-grade gliomas. Acta Neurol Belg 2004; 104(4): 148–153.

2. Essig M, Anzalone N, Combs SE, Dörfler À, Lee SK, Picozzi P, Rovira A, Weller M and Law M. MR Imaging of neoplastic central nervous system lesions: review and recommendations for current practice. Am J Neuroradiol 2012: 33(5): 803–817.

3. Bian W, Khayal IS, Lupo JM, McGue C, Vandenberg S, Lamborn KR, Chang SM, Cha S and Nelson SJ. Multiparametric characterization of grade 2 glioma subtypes using magnetic resonance spectroscopic, perfusion, and diffusion imaging. Transl Oncol 2009; 2(4): 271–280.
4. Bulakbasi N, Kocaoglu M, Farzaliyev A, Tayfun C, Ucoz T and Somuncu I. Assessment of diagnostic accuracy of perfusion MR imaging in primary and metastatic solitary malignant brain tumors, Am J Neuroradiol 2005; 26(9): 2187–2199.
5. Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, Melhem ER and Davatzikos C. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med 2009; 62(6): 1609–1618.
6. Cha S. Update on brain tumor imaging: from anatomy to physiology. Am J Neuroradiol 2006; 27(3): 475–487.
7. Jude Hemanth D, Kezi Selva Vijila C and Anitha J. Fuzzy based experimental verification of significance of skull tissue removal in brain tumor image segmentation. International Journal of Computer Applications 2010; 1(29): 56–61.
8. Lim TS, Iaria G, Moon SY (2010) Topographical disorientation in mild cognitive impairment: a voxel-based morphometry study. J Clin Neurol 6: 204–211.
9. Leube DT, Weis S, Freymann K, Erb M, Jessen F, et al. (2008) Neural correlates of verbal episodic memory in patients with MCI and Alzheimer’s disease—a VBM study. International Journal of Geriatric Psychiatry 23: 1114–1118.
10. Beeson PM, King RM, Bonakdarpour B, Henry ML, Cho H, et al. (2011) Positive Effects of Language Treatment for the Logopenic Variant of Primary Progressive Aphasia. J Mol Neurosci.
11. Heinzel A, Minnerop M, Schafer R, Muller HW, Franz M, etal. (2011) Alexithymia in healthy young men: A voxel-based morphometric study. J Affect Disord.
12. Tavanti M, Battaglini M, Borgogni F, Bossini L, Calossi S, et al. (2011) Evidence of diffuse damage in frontal and occipital cortex in the brain of patients with post-traumatic stress disorder. Neurol Sci.
13. Beste C, Saft C, Konrad C, Andrich J, Habbel A, et al. (2008) Levels of error processing in Huntington’s disease: a combined study using event-related potentials and voxel-based morphometry. Hum Brain Mapp 29: 121–130.
14. Ille R, Schafer A, Scharmuller W, Enzinger C, Schoggl H, et al. (2011) Emotion recognition and experience in Huntington disease: a voxel-based morphometry study. J Psychiatry Neurosci 36: 100143.
15. Zou K, Deng W, Li T, Zhang B, Jiang L, et al. (2010) Changes of brain morphometry in first-episode, drug-naive, non-late-life adult patients with major depression: an optimized voxel-based morphometry study. Biol Psychiatry 67: 186–188.
16. Brieber S, Neufang S, Bruning N, Kamp-Becker I, Remschmidt H, et al. (2007) Structural brain abnormalities in adolescents with autism spectrum disorder and patients with attention deficit/hyperactivity disorder. J Child Psychol Psychiatry 48: 1251–1258.