Anatomical Brain MRI Segmentation Methods: Volumetric Assessment of the Hippocampus

Flávio Luiz Seixas
Débora Christina Muchaluat Saade
Aura Conci
Computer Science Institute,
Federal Fluminense University,
Niterói – RJ, Brazil
{fseixas, debora, aconci}@ic.uff.br

Andrea Silveira de Souza
Fernanda Tovar-Moll
Ivanei Bramatti
LABS D´Or Network,
Radiology Department,
Rio de Janeiro – RJ, Brazil
{andrea.silveiradesouza, tovarmollf,
ibramati}@gmail.com

Abstract—Alzheimer's disease (AD) is the most common cause of dementia among elderly, initially taking place in the hippocampus. Volumetric evaluation of the hippocampus can be an important biomarker for AD, but its manual volumetric assessment is time-consuming and depends on previous and specific anatomical knowledge, forbidding its use on clinical practice. This work analyzes three medical image analyzing tools, Freesurfer, IBASPM and FSL, in order to identify the best method for hippocampus volumetric assessment. Nineteen magnetic resonance image (MRI) from cognitively intact individuals were used for qualitative comparison among these methods. More concordant results were found using FSL and Freesurfer in terms of hippocampus delineation.

Keywords- automated image analysis methods, magnetic resonance imaging, Alzheimer's disease, hippocampus, volumetric segmentation.

I. INTRODUCTION

Alzheimer's disease (AD) is the most common cause of dementia (49.9% to 84.5% of all dementia in Latin America), with a progressive course, beginning with neuronal dysfunction, and evolving to irreversible loss of neurons [1]. Early AD detection and treatment allow better clinical results, preserving the loss of cognitive functions for longer periods, with great social-economical impact. Subtle hippocampus and amygdala volumetric atrophy can be observed even in early AD, becoming an important biomarker for the disease, especially with disease progression. Volumetric MRI studies can already indicate small volume loss of those anatomical structures [2]. However, hippocampus manual 3D identification relies mainly on highly time-consuming segmentation procedures. An expert may need 30 minutes to trace a single structure such as the hippocampus [3]. Moreover, automated methods have provided consistent results comparable to manual segmentation [4].

This paper compares three fully automated methods implemented on public-domain applications: (1) Freesurfer, (2) IBASPM and (3) FIRST/FSL. Freesurfer (http://surfer.nmr.mgh.harvard.edu/) is a set of software tools for the study of cortical and subcortical anatomy developed by members of Athinoula A. Martinos Center for Biomedical Imaging [5]. IBASPM (http://www.thomaskoenig.ch/Lester/ibaspm.htm) is a toolbox for brain segmentation of structural MRI, developed by Cuban-Neuroscience Center [6]. FIRST/FSL (http://www.fmrib.ox.ac.uk/fsl) is a model-based segmentation and registration tool applied to get subcortical brain segmentation using Bayesian shape and appearance models, developed by FMRIB group from Oxford University [7]. Our primary goals are: (1) analyze and synthesize the imaging processing for each of the three tools above described; (2) compare the left/right hippocampus volumes obtained in each method, quantifying the volume agreement among them; and (3) discuss the quality of the volume segmentation obtained with each method.

Regarding related works, Powell *et al.* [8] compared four fully automated brain structure segmentation methods: template-based, probabilistic-based, artificial neural network-based (ANN) and support vector machine-based (SVM) segmentation. The image base was composed by 25 normal control subjects. The segmentation methods results were compared to manually defined regions. The comparison evaluation between automatic and manual segmentation was performed using three overlapping metrics: relative, similarity and spatial. The relative overlap for the hippocampus was lower than 59%. The comparison evaluation between automatic segmentation was performed using *t-test p-values* of relative overlap measured between the segmentation methods. The results showed a little disparity between the ANN and SVM, getting a *p-value* upper to 0.48. Morey *et al.* [9] compared the performance of FSL/FIRST and Freesurfer. The left/right hippocampus volumes derived from each automated measurement were compared to expert hand tracing for percent volume overlap, also called Dice's coefficient, and across-sample correlation. A total of 20 normal control subjects were collected. When comparing automated measures to manual tracing, the left-right hippocampus segmentation volume overlap produced by Freesurfer (83% and 82% respectively) was superior to FIRST (79% and 80% respectively). The conclusions were that the automated methods generated systematically larger volumes than manual tracing. Freesurfer and FIRST are not equal when compared to manual tracing and Freesurfer was superior for

segmenting the hippocampus by all the objective measurement performed. Tae *et al.* [10] included T1-weighted MRI of 21 patients with chronic major depressive disorder (MDD) and 20 normal controls. They compared the hippocampus volume achieved from two fully automated tools to manual segmentations. The automated segmentation methods use surface-based using Freesurfer and individual-atlas using IBASPM. The agreement test between the volumetric methods got an ICC right 0.84 and left 0.85 when comparing Freesurfer to manual, and ICC right 0.65 and left 0.71 when comparing IBASPM to manual. Their conclusions were that the automated hippocampal volumetric methods showed good agreement with manual hippocampal volumetry, the absolute volume measured using Freesurfer was 35% larger than manual and the agreement was questionable with IBASPM.

Different from our work, those related works did not perform the comparison among different automated-segmentation methods. They compared automated segmentation with manual segmentation. Manual segmentation is sensible to manual tracing method and expert's skill, as well as it is highly time-consuming [4].

II. AUTOMATED IMAGE ANALYSIS METHODS

A. *Image Database*

MR images used in this work were provided by LABS D´Or Network Research Group, Rio de Janeiro – Brazil. Nineteen exams from distinct cognitively intact patients, scanned on high field Philips magnets, were divided in two groups: group I, aged from 37 to 57 years old (average of 48), 4 male and 9 female, was scanned in a 1.5 Tesla MR unit, using images with 256x256 resolution and a total of 128 slices; group II, aged from 21 to 54 years old (average of 33), 1 male and 8 female, was scanned in a 3 Tesla MR unit (total of 9 exams), using images with 512x512 resolution and a total of 386 slices. Volumetric T1-weighted images were acquired in the sagittal plane.

B. *Hippocampus Anatomical Definition*

The hippocampus is part of middle arc of the limbic system. It is located in the medial temporal lobe inferior to the choroidal fissure and temporal horn. The gray matter of the hippocampus is an extension of the subiculum of the parahippocampal gyrus. The hippocampus itself consists of two interlocking C-shaped structures: the cornu of Ammonis and the dentate gyrus [11]. Figure 1 shows the hippocampus representation extracted from Purves [12].

C. *Automated Image Segmentation Methods*

Medical image segmentation methods can be grouped into two classes: (1) voxel-based methods and (2) vertex-based methods. Voxel-based methods are region-based segmentation methods, where the anatomical structures are represented by their voxel features (signal intensity, spatial position) and neighborhood-based features. Vertex-based methods are boundary-based segmentation methods, where the anatomical structures are represented by their boundary surface parameterized by vertices grouped in a mesh.

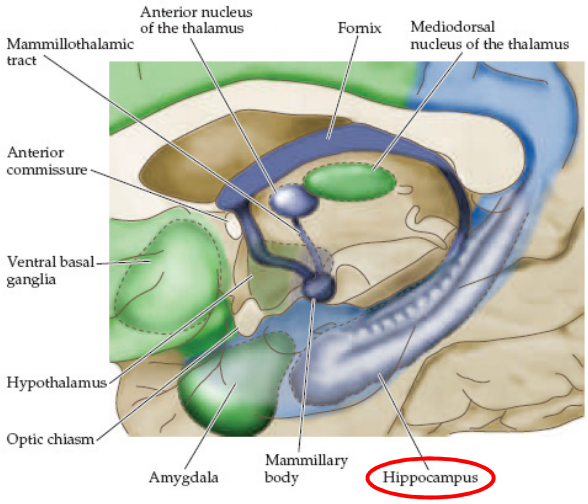

Figure 1. Subcortical structures graphical representation

Figure 2 shows the general image segmentation process used by the three methods applied on this work. In the preprocessing phase, we performed image reorientation and image format conversion to NIFTI files (http://nifti.nimh.nih.gov/). Despite the axial slicing information being saved in image header section, we preferred to deal with the source image with the same axial slicing as the standard template image: Right-Left, Posterior-Anterior and Inferior-Superior, based on radiological convention orientation. Image registration is the process of aligning images so that corresponding features can easily be related [13]. The image registration stage was used by the three studied methods to align the source image to a common space, achieving a voxel-to-voxel correspondence for all subjects. The goal of the segmentation stage is to simplify the representation of an image into something that is more meaningful and easier to analyze. In the segmentation stage, the brain tissue is segmented into four or five classes: gray matter, white matter, cerebro-spinal fluid, non-brain regions and skull. In the labeling stage, the brain structures mapped into an anatomical atlas are assigned to the source image. Finally, the volume stage computes the volume of the structures of interest.

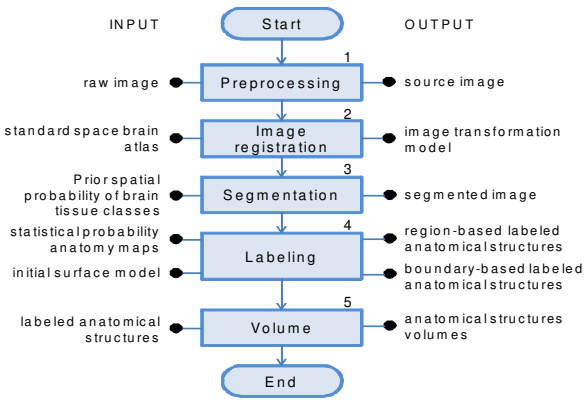

Figure 2. General image segmentation process

Table I summarizes the main differences among the three methods studied, grouping them into the generic image segmentation defined previously.

TABLE I. IMAGE SEGMENTATION PROCESS COMPARISON

Freesurfer	Image registration	Affine registration driven by 9 degrees of freedom.
	Template image	Tailarach space.
	Segmentation	Removes the skull scalp.
	Labeling	Generated by assigning each voxel to the class for which the probability is the greatest.
	Anatomical atlas	GCA atlas (each structure is represented by a class in a probability function).
	Volume	It counts the labeled structures assigned in the previous stage.
FSL	Image registration	Performed in two stages. The first makes a standard 12 degrees of freedom registration to the template. The second stage applies a 12 degrees of freedom registration using MNI 152 subcortical mask to exclude voxels outside the subcortical regions.
	Template image	MNI 152 space at 1mm resolution.
	Labeling	It performs a shape and appearance active surface model.
	Anatomical atlas	Shape appearance models constructed from 336 subjects, consisting of children and adults, normal individuals and subjects with pathologies. The result is a vertex set structured in a mesh.
	Volume	Volume is computed filling the meshes obtained from the previous stage.
IBASPM	Image registration	The image registration and segmentation image segmentation stages are performed in a unified model. The registration is driven by 12 degrees of freedom and a non-linear registration model. The result is a transformation composed by an affine matrix and non-linear registration model.
	Template image	MNI ICBM AVG 152 smoothing with 8mm Gaussian kernel.
	Segmentation	The input parameters are the brain tissue probability maps published by ICBM (International Consortium of Brain Mapping). It performs an iterative probability function maximization using a Bayesian framework.
	Labeling	It uses the gray matter mask segmented previously. The brain structures presented in the atlas are mapped to the source image using a simple binary association.
	Anatomical atlas	AAL atlas (binary atlas)
	Volume	The source image is deformed again applying the inverse transformation model and labeled structures are carried to the original space.

III. SEGMENTATION RESULTS AND EVALUTATION

With the purpose of quantify the volume agreement of the structures segmented automatically by the three methods studied, we propose two volume agreement indexes: (1) the relative volume overlay *(RVO)*, as shown in Equation 1, and (2) the intraclass correlation *(ρ),* as shown in Equation 2 .

$$RVO(A_1;A_2;...A_i) = \frac{A_1 \cap A_2 \cap ... A_i}{A_1 \cup A_2 \cup ... A_i} \cdot 100 \quad (1)$$

where A_1, A_2, ...A_i are voxel sets belonging to the structure of interest from the i^{th} medical image analyzing package, respectively.

$$\rho(A;B) = \frac{E[(A-\mu_A)\cdot(B-\mu_B)]}{\sigma_A \cdot \sigma_B} \quad (2)$$

where ρ is the correlation between variables *A* and *B, E[]* is the expected value function and μ_i, and σ_i are the mean and variance of variable *i*, respectively.

In order to work with average volumes, it is necessary to use normalized volumes. A common denominator usually selected for this purpose is the intracranial volume. We used the estimated total intracranial volume *(eTIV)* obtained from Freesurfer [14]. Equation 3 shows how the normalized volume for each individual is obtained. The mean normalized volumes shown in Table II were obtained by corresponding arithmetic average. Table III shows the relative overlay volumes *(RVO)* and the intraclass correlation index *(ρ)*. The volumes are correlated when the index is close to one (or minus one).

$$nVol_i = \frac{Vol_i}{eTIV_i} \cdot 10000 \quad (3)$$

where Vol_i is the structure volume of a given individual and $eTIV_i$ is the estimated total intracranial volume.

TABLE II. AVERAGE NORMALIZED VOLUMES

Hipp.[c]	Group[b]	Freesurfer[a]	IBASPM[a]	FSL[a]
Left (LH)	I	29.9 (2.7)	20.1 (1.7)	34.8 (3.5)
	II	27.1 (1.3)	22.6 (3.2)	29.5 (1.6)
Right (RH)	I	29.1 (4.0)	15.9 (2.3)	35.4 (5.5)
	II	26.3 (1.8)	17.3 (1.7)	29.1 (2.2)

a. Volume normalized by estimated total intracranial volume as shown in Equation 3 represented by its average and standard deviation between parentheses.
b. Group I is composed by exams scanned in a 1.5 Tesla MR unit (total of 10) and Group II is composed by exams scanned in a 3 Tesla MR unit (total of 9).
c. We have abbreviated hippocampus as Hipp.

TABLE III. RELATIVE OVERLAY VOLUMES *(RVO)* AND INTRACLASS CORRELATION INDEX *(ρ)*

		Freesurfer IBASPM		Freesurfer FSL		IBASPM FSL		Freesurfer IBASPM FSL
		RVO[a]	ρ[b]	***RVO***[a]	ρ[b]	***RVO***[a]	ρ[b]	***RVO***[a]
LH	I	43.00	0.76	61.75	0.60	61.42	0.43	30.72
	II	45.35	0.49	67.59	0.92	60.95	0.43	35.42
RH	I	23.23	0.32	61.48	0.76	56.26	0.30	14.85
	II	20.35	0.81	63.61	0.70	52.92	0.54	12.33

a. Relative overlay volume average. The unit is percentage (%)
b. Intraclass correlation index. An index equal to one (or minus one) means that both volumes are correlated. An index equal to zero means that the correlation does not exist.

Freesurfer and FSL presented better relative overlay indexes as shown in Table III. Regarding the intraclass correlation, Freesurfer and IBASPM presented better results in general, as shown in Table III. Despite referring to the same anatomical structure, the three methods diverged when considering relative overlay volume. The highest relative overlay volume value was about 35%. Figure 3 shows the segmentation results of three computational methods used in this study, displayed on 2D coronal images. When the sulcus between the hippocampus and the parahippocampus gyrus is not clearly defined, the upper cortex of the parahyppocampus gyrus was included in the hippocampus volume segmentation, as shown in Figure 3 (a1, b1, c1). Probably, this problem would not occur in

the elderly population as this sulcus would be more pronounced between these two structures, allowing better boundaries discrimination. Regarding Freesurfer, the dentate gyrus, subiculum and upper area of parahippocampal gyrus (entorhinal cortex) appear to have been included in hippocampus volume, as shown by the red arrows in Figure 3 (a3). Freesurfer included also the inferior cornu of the lateral ventricle. In the most posterior images, Freesurfer included part of the fornix, as shown by the red arrow in Figure 3 (a1). The Freesurfer definition of the hippocampus anatomical boundaries is not clearly presented in program's technical manual [10]. We noticed that the hippocampal segmentation obtained from IBASPM was discontinuous with clear exclusion of easily recognizable areas belonging to the hippocampus, as shown in Figure 3 (b3, b4). The errors in hippocampal volume segmentation from IBASPM arose from inaccurate image registration and use of the MNI single-subject for the manually predefined region of interest of the hippocampus [15]. In FSL, the limits of the hippocampus have been expanded, including the adjacent white matter and other neighbor structures, providing volumes significantly inflated. FSL also included the fornix and entorhinal cortices.

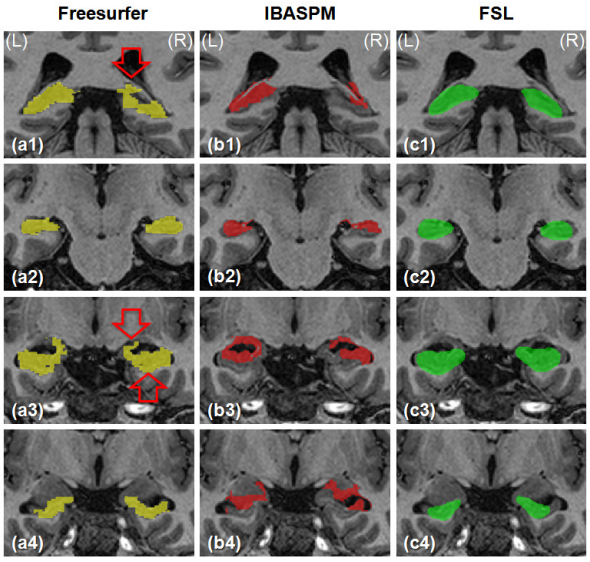

Figure 3. Coronal right (R) and left (L) hippocampus slices. segmented with Freesurfer, FSL and IBASPM: Labels from 1 to4 indicate the most posteriorto the most anterior slices, respectively.

IV. FINAL REMARKS

An important aspect of automated volumetric methods is its reproducibility, assuring consistent findings. However, accuracy is very important for volumetric studies. We provided a complete and detailed comparison of three medical imaging packages applied to get hippocampus volumetric measures. Volume measures show disagreements among the three methods, indicating the need of more exhaustive research in this field. Despite being sensible to human skills and time-consuming, manual segmentation is used as a reference from which might be calculated a deviation. As future works, we will perform comparisons against manual segmentation, in order to evaluate the hippocampus segmentation accuracy.

ACKNOWLEDGMENTS

This work was partially funded by FAPERJ, CNPq and CAPES.

REFERENCES

[1] R. Nitrini, C. M. C. Bottino, C. Albala, N. S. C. Capuñay, C. Ketzoian, J. J. L. Rodriguez, G. E. Maestre, A. T. A. Ramos-Cerqueira, and P. Caramelli, "Prevalence of dementia in Latin America: a collaborative study of population-based cohorts," International Psychogeriatrics, vol. 21, pp. 622-630, 2009.

[2] K. Kantarci and C. R. Jack, "Neuroimaging in Alzheimer's disease: an evidence-based review," Neuroimaging Clinics of North America, vol. 13, pp. 197-209, 2003.

[3] M. J. Firbank, R. Barber, E. J. Burton, and J. T. O'Brien, "Validation of a fully automated hippocampal segmentation method on patients with dementia," Human Brain Mapping, vol. 29, pp. 1442-1449, 2008.

[4] S. Klöppel, C. M. Stonnington, J. Barnes, F. Chen, C. Chu, C. D. Good, I. Mader, L. A. Mitchell, A. C. Patel, C. C. Roberts, N. C. Fox, C. R. Jack, J. Ashburner, and R. S. J. Frackiowiak, "Accuracy of dementia diagnosis: a direct comparison between radiologists and a computerized method," Brain, vol. 131, pp. 2969-2974, 2008.

[5] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. v. d. Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. M. Dale, "Whole brain segmentation automated labeling of neuroanatomical structures in the human brain," Neurotechnique, vol. 33, pp. 341-355, 2002.

[6] Y. Alemán-Gómez, L. Melie-Garcia, and P. Valdés-Hernandez, "IBASPM: Toolbox for automatic parcellation of brain structures," in Meeting of the Organization for Human Brain Mapping. vol. 27 Florence: Neuroimage, 2006.

[7] B. Patenaude, "Bayesian statistical models of shape and appearance for subcortical brain segmentation," in Oxford Centre for Functional Magnetic Resonance Imaging of the Brain. vol. Doctorate: University of Oxford, 2007.

[8] S. Powell, V. A. Magnotta, H. Johnson, V. K. Jammalamadaka, R. Pierson, and N. C. Andreasen, "Registration and machine learning-based automated segmentation of subcortical and cerebellar brain structures," Neuroimage, vol. 39, pp. 238-247, 2008.

[9] R. A. Morey, C. M. Petty, Y. Xu, J. P. Hayes, H. R. Wagner, D. V. Lewis, K. S. Labar, M. Styner, and G. McCarthy, "A comparison of automated segmentation and manual tracing for quantifying of hippocampal and amygdala volumes," Neuroimage, vol. in press, 2009.

[10] W. S. Tae, S. S. Kim, K. U. Lee, E.-C. Nam, and K. W. Kim, "Validation of hippocampal volumes measured with one manual and two automated methods using Freesurfer and IBASPM in chronic major depressive disorder," Neuroradiology, vol. 51, pp. 203-204, 2009.

[11] L. P. Mark, D. L. Daniels, T. P. Naidich, Z. Yetkin, and J. A. Borne, "The hippocampus," American Journal of Neuroradiology, vol. 14, pp. 709-712, 1993.

[12] D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A.-S. Lamantia, J. O. McNamara, and S. M. Williams, Neuroscience, Third ed. Sunderland: Sinauer Associates, 2004.

[13] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, Medical image registration vol. 1: CRC Press LLC, 2001.

[14] R. L. Buckner, D. Head, J. Parker, A. F. Fotenos, D. Marcus, J. C. Morris, and A. Z. Snyder, "A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume," Neuroimage, vol. 23, pp. 724-738, 2004.

[15] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," Neuroimage, vol. 15, pp. 273-289, 2002.