Computers in Biology and Medicine 95 (2018) 90–98

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Atlas selection for hippocampus segmentation: Relevance evaluation of three meta-information parameters☆

Vanderson Dill [a], Pedro Costa Klein [a], Alexandre Rosa Franco [a,b,c], Márcio Sarroglia Pinho [a,*]

[a] School of Technology, PUCRS, Porto Alegre, Brazil
[b] School of Medicine, PUCRS, Porto Alegre, Brazil
[c] Brain Institute of Rio Grande do Sul, PUCRS, Porto Alegre, Brazil

ARTICLE INFO

Keywords:
Hippocampus segmentation
Atlas-based segmentation
Magnetic resonance imaging
Medical imaging
Segmentation evaluation
Alzheimer's disease

ABSTRACT

Current state-of-the-art methods for whole and subfield hippocampus segmentation use pre-segmented templates, also known as atlases, in the pre-processing stages. Typically, the input image is registered to the template, which provides prior information for the segmentation process. Using a single standard atlas increases the difficulty in dealing with individuals who have a brain anatomy that is morphologically different from the atlas, especially in older brains. To increase the segmentation precision in these cases, without any manual intervention, multiple atlases can be used. However, registration to many templates leads to a high computational cost. Researchers have proposed to use an atlas pre-selection technique based on meta-information followed by the selection of an atlas based on image similarity. Unfortunately, this method also presents a high computational cost due to the image-similarity process. Thus, it is desirable to pre-select a smaller number of atlases as long as this does not impact on the segmentation quality. To pick out an atlas that provides the best registration, we evaluate the use of three meta-information parameters (medical condition, age range, and gender) to choose the atlas. In this work, 24 atlases were defined and each is based on the combination of the three meta-information parameters. These atlases were used to segment 352 vol from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Hippocampus segmentation with each of these atlases was evaluated and compared to reference segmentations of the hippocampus, which are available from ADNI. The use of atlas selection by meta-information led to a significant gain in the Dice similarity coefficient, which reached 0.68 ± 0.11, compared to 0.62 ± 0.12 when using only the standard MNI152 atlas. Statistical analysis showed that the three meta-information parameters provided a significant improvement in the segmentation accuracy.

1. Introduction

The analysis of morphometric characteristics of the hippocampus or its subfields is an important process in the diagnosis of many neurological and neuropsychological diseases, including temporal lobe epilepsy [1], Alzheimer's disease (AD) and mild cognitive impairment (MCI) [2], schizophrenia [3], major depression [4], bipolar disorder [5], and post-traumatic stress syndrome [6], among others [7].

Manually segmenting the hippocampus and calculating its volume are laborious tasks and are prone to subjective interpretation by health professionals. Automated methods that can reduce the subjectivity and increase the segmentation accuracy are highly desirable. However, automatic segmentation of the hippocampus in magnetic resonance images (MRI) presents some challenges. In a T1-weighted image, the pixel (or voxel) intensities of the hippocampus are similar to the intensities from other nearby brain structures, such as the amygdala, caudate nucleus, and thalamus [8]. Also, well-defined borders of the hippocampus with its neighboring structures are not easily identifiable, partial volume effects make pixel classification at the border more difficult, and the non-uniformity of intensities hinders the process of segmenting the hippocampus [9].

When performing a manual segmentation, the radiologist benefits

☆ Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators from ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

* Corresponding author. School of Technology, PUCRS – Pontifícia Universidade Católica do Rio Grande do Sul, Building 32, Office 607, Av. Ipiranga 6690, CEP: 90619-000, Porto Alegre, RS, Brazil.
E-mail address: pinho@pucrs.br (M.S. Pinho).

https://doi.org/10.1016/j.compbiomed.2018.02.005
Received 13 November 2017; Received in revised form 7 February 2018; Accepted 8 February 2018

from his or her previous knowledge, such as the position of the hippocampus in the brain, the relative positions of neighboring structures, and knowledge about the usual shape of the hippocampus, which allows the professional to overcome those limitations. In automated methods, these high-level features cannot be obtained directly from pixel intensities. However, these features can be incorporated by other means, such as by the usage of atlases [10]. An atlas is an image in which there is reference segmentation for structures of interest, obtained through manual or automatic methods.

The simplest technique of atlas usage for hippocampus segmentation consists of registering a single atlas to the target image. With the deformation map acquired through this registration, it is possible to warp the existing segmentation in the atlas to the target image to obtain the final segmentation. The quality of the segmentation achieved through the **single-atlas** based method is strongly dependent on the choice of atlas and the registration accuracy. To obtain acceptable results, the method must use an atlas that has been created from individuals with anatomies similar to that of the individual in the target image, since the available registration techniques cannot align individuals with large anatomical differences with the required precision [11].

As a way to overcome this problem, many current methods use **multi-atlas** based techniques. Several different techniques have been used for applying multiple atlases in hippocampus segmentation. One of the most common ways is to individually segment the target with all the available atlases. Then, the creation of the final segmentation is done through label fusion techniques [12]. Many methods have been presented within the last years in this field of research [10,13–27].[1] However, despite showing good accuracy, this kind of segmentation method based on multiple atlases has a high computational cost, since the target image must be registered separately with each atlas used. Consequently, it is desirable to use a smaller number of atlases as long as this does not impact on the segmentation quality.

For **atlas selection**, the most common methods are based on image similarity and patient meta-information. Selection by image similarity gives excellent results but still has a high computational cost. On the other hand, selection by meta-information has the advantage of having a significantly low computational cost but may not lead to the selection of an ideal atlas [28].

Thus, the usage of mixed approaches in which a pre-selection is made through meta-information and a selection by similarity is performed later for a reduced group of atlases may combine the best of each technique: the low computational cost of the meta-information selection and the effectiveness of the selection by image similarity techniques. Aljabar et al. [29], for example, selected a subset of a database with 275 atlases based on global characteristics of the image and meta-information, such as the age and sex of the patient. Local similarity metrics for the structure of interest can also be used [28,30–35]. Methods for segmenting other structures have also used similar approaches to select the ideal smallest number of atlases [36–38].

Taking the meta-information selection approach in isolation, the minimal number of atlases needed is equal to the product of the total number of possible values of each parameter used in the selection, that is, an atlas for each combination of the characteristics used. Therefore, knowing the relevance of selection parameters is important, since non-relevant parameters may be excluded from the selection process, thus reducing the total number of atlases needed.

In this context, the presented paper evaluates the relevance of three meta-information parameters in the selection of atlases for hippocampus segmentation in a dataset from the Alzheimer's Disease Neuroimaging Initiative (ADNI). For this purpose, we performed a large-scale experiment, in which 352 vol were segmented with 25 different atlases, (24 specific atlases plus the MNI152 atlas), representing different characteristic combinations of the following parameters: gender, age group, and clinical situation. Each target image and atlas pair was evaluated through the use of the DSC (Dice Similarity Coefficient) precision index [39] so that the influence of each of the three parameters on the result of the hippocampus segmentation could be evaluated.

[1] For a more complete reference, a review of hippocampal segmentation methods is presented in Dill et al. [52].

2. Materials and methods

The experiment was conducted on T1-weighted MRI data obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD).

The following sections describe the segmentation method used (Section 2.1), atlases (2.2), test data (2.3), evaluation procedure (2.4), and execution details (2.5).

2.1. Segmentation method

Fig. 1 shows a schematic overview of the applied segmentation method. Initially, the atlas selection is performed through meta-information matching. Using the parameters gender, clinical situation, and age group, an atlas that corresponds exactly to these three characteristics of the target volume is chosen as a template.

Details about the atlases used in this experiment are presented in Section 2.2. The method assumes the existence of a set generated in such a way that there is one atlas for each combination of possible values of parameters used (sex, age, and clinical condition). The values used for each selection parameter are presented in Table 1. Thus, to cover the full range of possible combinations of these values, 24 atlases are needed.

The target volume and the atlas are converted to the same coordinate space through a linear registration technique [40] using the MNI152 as the template [41]. This procedure is found to be necessary since it normalizes the images for future steps of skull stripping and allows the usage of non-linear registration techniques, which are designed for the fine alignment of internal structures and cannot handle large shifts.

After the linear registration, a skull stripping technique is applied on the target volume [42,43]. The BET (Brain Extraction Tool), which is part of the FSL toolkit (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/) [44], is used in this process. This execution step removes the skull from the image volume, keeping only the brain mass, thus increasing the non-linear registration precision which will be applied later. This also reduces the computational cost of the process, since a smaller number of pixels are used in the computation.

With the brain mass segmented, the non-linear registration algorithm is applied, aligning the atlas with the target volume. The result of the non-linear registration process is a deformation map that, when applied to the target image, aligns the brain as a whole, including the cerebral substructures. In Fig. 2, examples of these steps are presented. To acquire the final hippocampus segmentation, the linear and non-linear deformation maps are inverted and applied over the hippocampus segmentation from the atlas, generating a segmentation of the hippocampus of the target image.

For this experiment, two non-linear registration techniques were tested: **ART** (Automatic Registration Tool) [45,46], and **SyN** (Symmetric Diffeomorphic Image Registration) [47]. These techniques were mentioned in a previous study [48] as the most precise with regard to hippocampal region alignment. The SyN implementation used here is provided by the author through the ANTs tool (http://stnava.github.io/ANTs/) and ART was used through the tools available at the Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC, https://www.nitrc.org/projects/art).

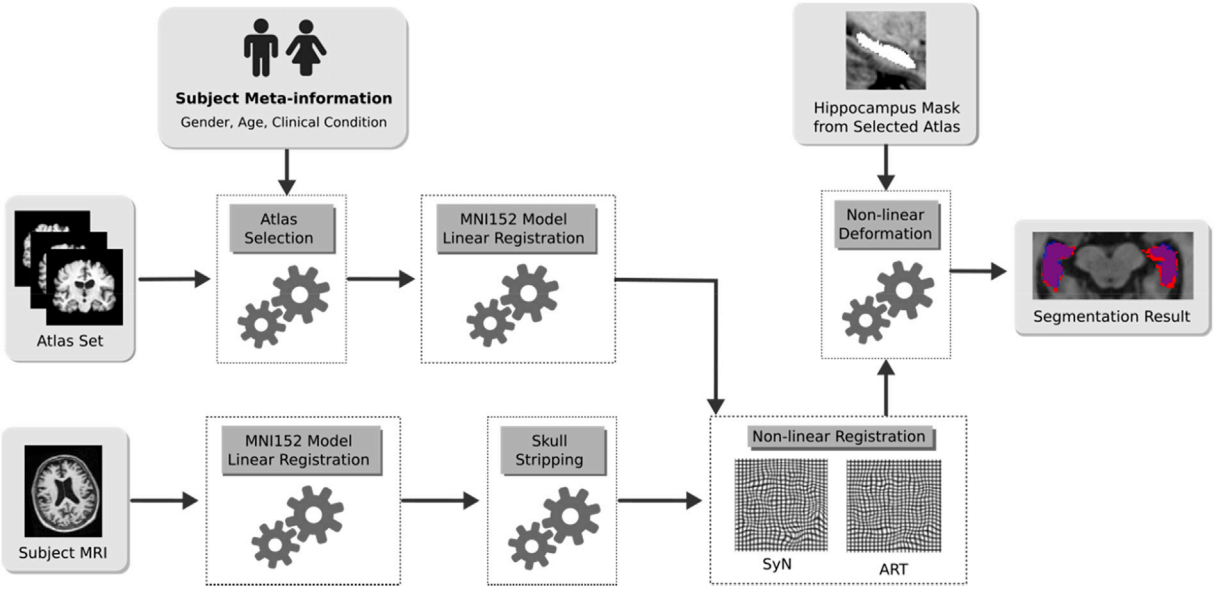

Fig. 1. Overview of the employed hippocampus segmentation method.

Table 1
Parameters used for atlas selection.

Parameter	Utilized values
Gender	Male (M), Female (F)
Age	70–75, 75–80, 80–85, 85–90 years
Clinical Situation	Normal (N), Mild Cognitive Impairment (MCI), Alzheimer (AD)

2.2. Atlases

For our application, an atlas is composed of two image volumes: a structural MRI of the brain and a binary image containing a map of the corresponding hippocampus. In total 25 atlases were used in the experiment performed, of which 24 were chosen to represent a combination of characteristics related to the three parameters: gender, age group, and clinical situation (Table 1) and the last was the MNI152 atlas. This atlas is provided by the Harvard Center for Morphometric Analysis, which has the reference segmentation for the MNI152 model with 21 subcortical structures, including the hippocampus.

The 24 atlases representing the characteristics defined by the three parameters were selected from the ADNI database. The available data include patients of various age groups, ethnicities, and genders. The volumes available in the ADNI database include reference segmentations for the hippocampus. This reference is generated by FreeSurfer and validated by health professionals, providing a good ground truth image. For each combination of characteristics among the three tested parameters, a representative individual was chosen from the database, for example, one atlas for a male aged 70–75 years with MCI, another atlas for a male aged 70–75 years with AD, and so on. The individual was randomly selected from the available images and the data were checked manually, taking into account the image integrity and the correctness of hippocampus segmentation.

2.3. Test data

The image set used for testing is composed of 352 T1-weighted MRI volumes selected from the ADNI database. Not all of the images available on ADNI have reference segmentation for the hippocampus; therefore, only a subset of the whole ADNI dataset was used. In addition, only one volume per patient was used. Series of examinations of the patient's disease evolution are common in this database; in these cases, only the first acquisition of the set was used. From the image set resulting from this selection, some images were randomly excluded in such a way that the remaining data were distributed in an approximately uniform way according to the three selection characteristics used: gender, age group, and clinical situation. Table 2 shows the number of subjects within each class.

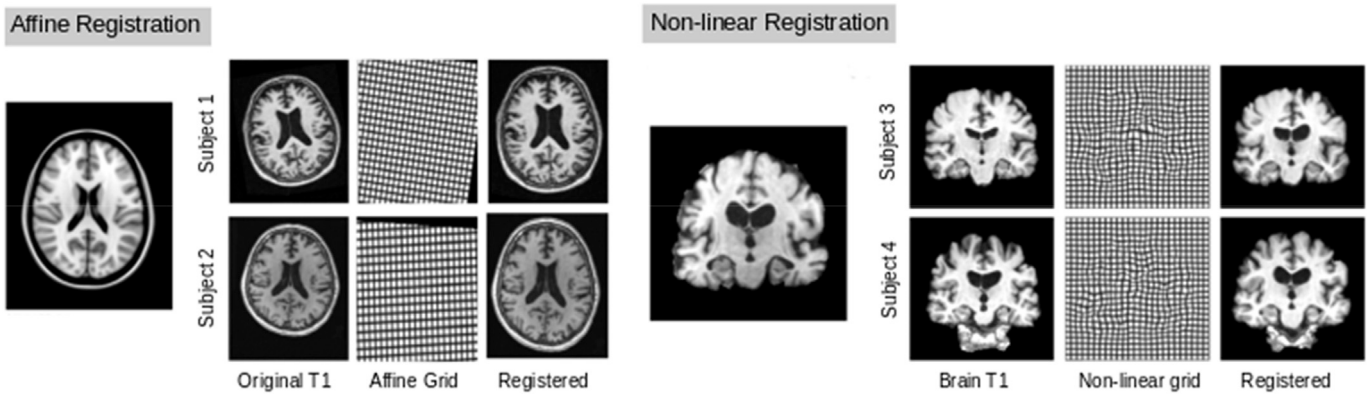

Fig. 2. Examples of deformation maps generated on the registration step.

2.4. Evaluation procedure

The evaluation of the influence of the atlas on segmentation accuracy aims to measure how this accuracy is affected when an atlas with the same characteristics as the individual being segmented (gender, age group, clinical situation) is used.

Each image of the testing set was segmented with each atlas extracted from the ADNI dataset and with the standard atlas MNI152 (see Section 2.2). Precision indexes obtained when using atlas selection by the meta-information method are compared with the indexes obtained using the **standard MNI152** atlas and using the average index obtained using a subset of the ADNI atlases.

This average index has the objective of evaluating the difference in precision between using an atlas selected according to the patient's characteristics – the **selected atlas** – or using a "randomly" selected atlas, as though no selection by meta-information were made at all. For the purpose of simplicity, from now on we will call this index the **average atlas** index.

To get the average atlas index, we compute the average DSC from a subgroup, as follows: In total, there are 24 parameter-specific atlases, one for each combination of meta-information, as stated in Section 2.2. So, we take a subgroup composed of the 23 remaining atlases, excluding the one that exactly matches the target characteristics, and compute the DSC from the average index obtained with each of them. For example, for the input group F; 75–80; AD, the index is the average of all the other specific atlases except the atlas F; 75–80; AD.

The experiment was run twice, employing SyN and ART registration techniques, over the same dataset.

2.5. Execution

In total, 17,600 executions were performed, in which every image from the 352 target images was segmented with each of the 25 atlases, using two versions of the algorithm (ART and SyN). The testing execution was parallelized in the High-Performance Laboratory of the Pontifical Catholic University of Rio Grande do Sul (LAD/PUCRS). This procedure was run on AMD Opteron 2.0-GHz CPUs and Intel Xeon 3.0-GHz bi-processor CPUs.

3. Results

Fig. 3 shows the segmentation results of three different individuals. This figure also presents slices of the T1 original image (upper image) and of the image with the generated hippocampus mask (lower image) resulting from the segmentation. In the segmented regions, the purple pixels represent the regions of the hippocampus in which there was agreement between the applied segmentation method and the segmentation from the ADNI database, which corresponds to the correct segmented region. The same image also displays pixels that only the ADNI database considered to be part of the hippocampus (blue pixels) and pixels that only the applied method considered to be part of (red pixels). The measurement of precision quantification is given by the DSC index, which represents the overlapping pixels in purple.

The best results were obtained by the method that uses the SyN registration technique with the atlas selected by meta-information (DSC index: 0.68 ± 0.11). The use of ART instead of SyN decreases the segmentation accuracy to 0.61 ± 0.10. The use of the MNI152 standard atlas instead of the selected atlas decreased the DSC index to 0.62 ± 0.12. The usage of the selected atlas also presented higher DSC values than the average atlas index, 0.64 ± 0.11 in this case.[2]

Table 2
Number of individuals in each category.

	70–75	75–80	80–85	85–90
F, AD	18	17	8	7
F, MCI	18	17	22	8
F, Normal	19	18	9	7
M, AD	17	12	14	7
M, MCI	16	18	22	19
M, Normal	20	18	14	7

The following sections present an analysis of the effects of selecting an atlas based on meta-information (3.1), the precision of the segmentation relative to age and the clinical situation (3.2), and the influence of each meta-information parameter used individually in the atlas selection (3.3). The results are evaluated statistically through paired t-tests, since we are comparing the DSC values of the same subjects but through different atlas-selection strategies. Since SyN performed better than ART with all 25 atlases, all the analyses that follow were performed on the results obtained using the SyN registration technique. All the resultant indexes obtained with both registration techniques are presented in the Supplementary Material.

3.1. Effects of atlas selection

The graphic in Fig. 4 presents the average of the precision indexes obtained by using the proposed method and the other two segmentation methodologies. For presentation purposes, the data are grouped initially by clinical situation of the patient (normal, MCI, or AD), then by the age group, and finally by gender (female or male). In 17 of the 24 groups, there was an increase in the average precision of the segmentation using the atlas selected according to the patient's condition. This happened both when comparing the proposed method to the usage of the MNI152 standard atlas and when comparing to the method that used the average atlas index.

The atlas selection is better for 21 groups when compared to the MNI Atlas, and similar indexes are obtained for two groups (M; 85–90; N and F; 85–90; A). The use of the MNI152 atlas was only better for one group (F; 70–75; N). Compared to the average atlas index, the use of the selected atlas has improved the precision index for 17 groups; for five of those the indexes are similar and for two the indexes are lower.

The average DSC index for all categories shows that atlas selection increased precision from 0.62 ± 0.12 to 0.68 ± 0.11 with MNI152 and the selected atlases, respectively. A paired t-test showed that this increase is statistically significant ($p < 0.05$; $x = 0.62$; $\mu_0 = 0.68$; $s = 0.11$; $n = 352$).

On comparing the general DSC obtained by the average index with selected atlas, it can be seen that the DSC index increased from 0.64 ± 0.11 to 0.68 ± 0.11. Again, a paired t-test confirmed that the increase in precision for the meta-information atlas is statistically significant ($p < 0.05$; $x = 0.64$; $\mu_0 = 0.68$; $s = 0.11$; $n = 352$).

3.2. Segmentation precision for age and progression of clinical situation

The graph in Fig. 5 shows the DSC precision based on the progression of the clinical situation. Analyzing the right side of the graph, where the MCI and AD patients are concentrated, it can be noticed that the use of both a standard MNI152 atlas and the average atlas index results in a clear precision loss, especially for patients with more advanced clinical situations.

This precision loss can be observed more easily by applying a linear regression to the indexes of each atlas. By observing the bars that represent the precision of each method relative to the progress of the characteristics used, it is noticeable that the precision level of the method that implements the meta-information atlas selection is practically constant, while in the other two methods, the precision level decreases proportionally to the worsening of the patient's clinical condition.

This behavior can be explained by the fact that in advanced cases of

[2] The DSC index obtained for each group of individuals, applying each tested atlas, is presented in the Supplementary Material.

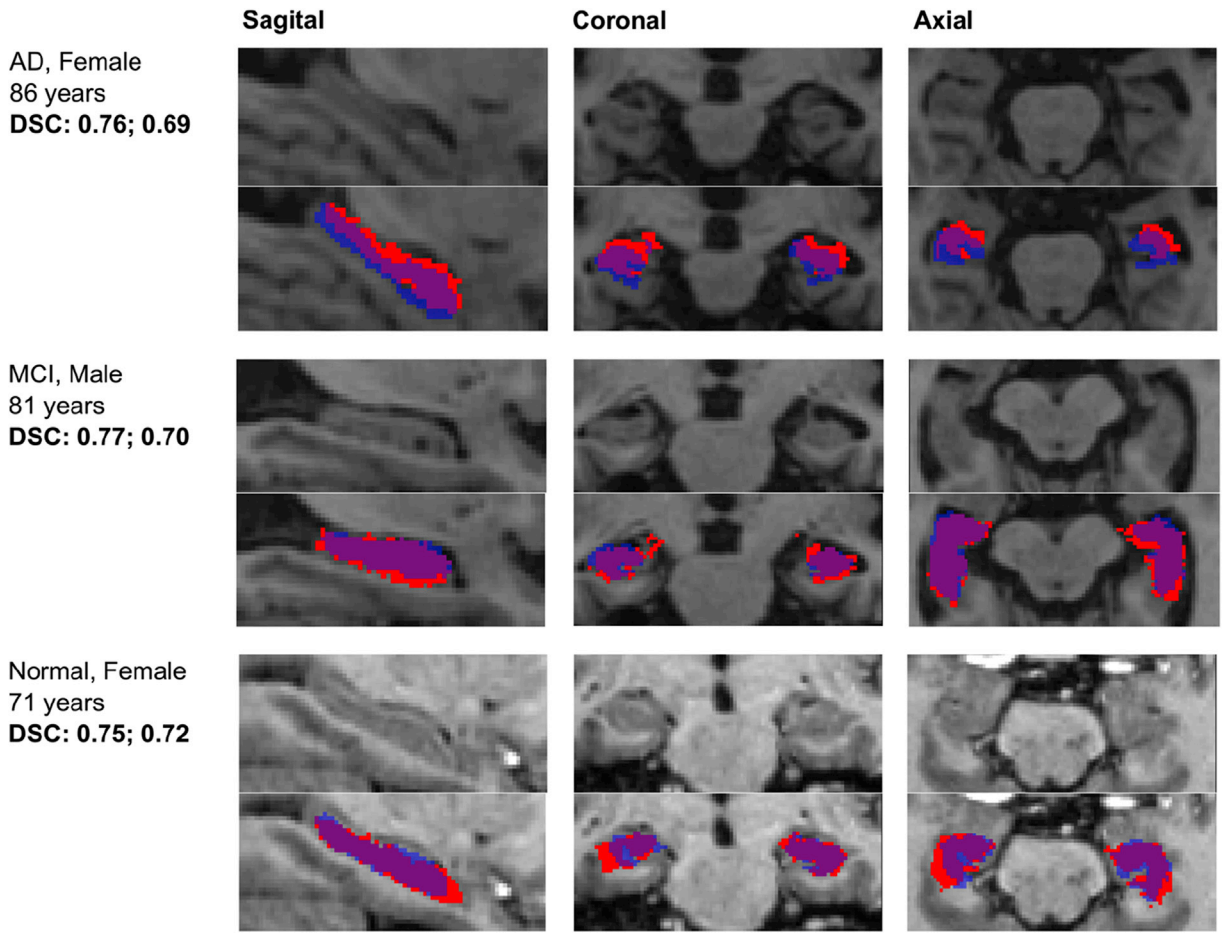

Fig. 3. Results of applying the proposed segmentation method. In red, the segmentation obtained with the applied method; in blue, the reference segmentation obtained from the ADNI; in purple, the overlap of both segmentations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

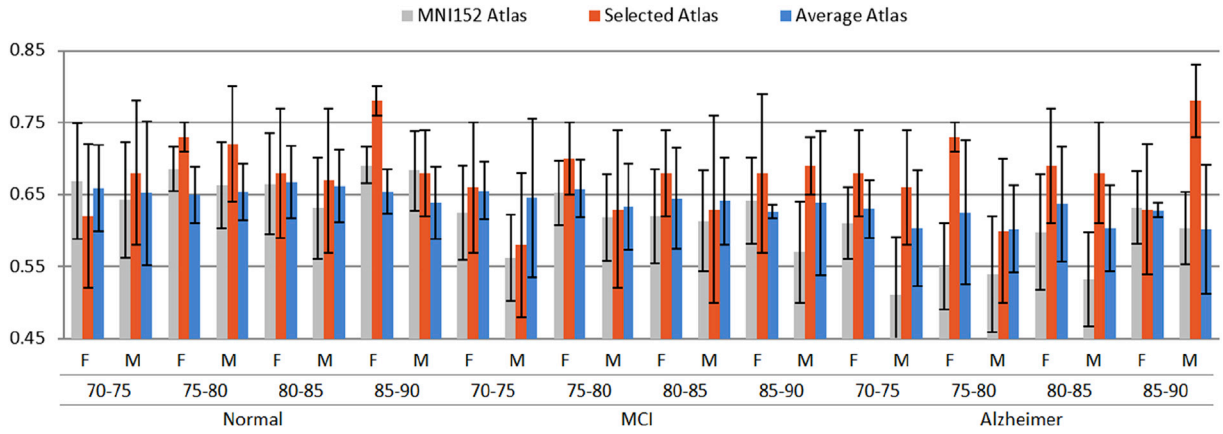

Fig. 4. Graphic comparing the DSC precision indexes obtained through the use of different atlases. Error bars represent standard deviation.

AD, the hippocampal atrophy is greater [49], leading to a greater difference from a normal hippocampus in terms of both size and shape, which makes the registration process more complex. Hence, the usage of selected atlases reduces this effect, since similar images demand less from the registration technique.

Regarding the age of the individuals, a small degradation of the precision indexes can be observed for older individuals when using the average atlas (Fig. 6), but it is less pronounced than in the cases of the clinical situation. This is not observed when using the atlas MNI152 even though the indexes are generally lower for this atlas. In relation to the clinical situation, it is perceivable that the loss of precision of the MNI152 atlas for unhealthy individuals is greater than that obtained when using the average atlas. The MNI152 Harvard-Oxford atlas is more "distant" from the tested individuals relating to the patient's age than the atlas built from the average of the ADNI database atlases, since the MNI is built from young healthy individuals, while the atlases used are built from individuals with three different clinical situations (normal, MCI, and AD), all in more advanced age groups.

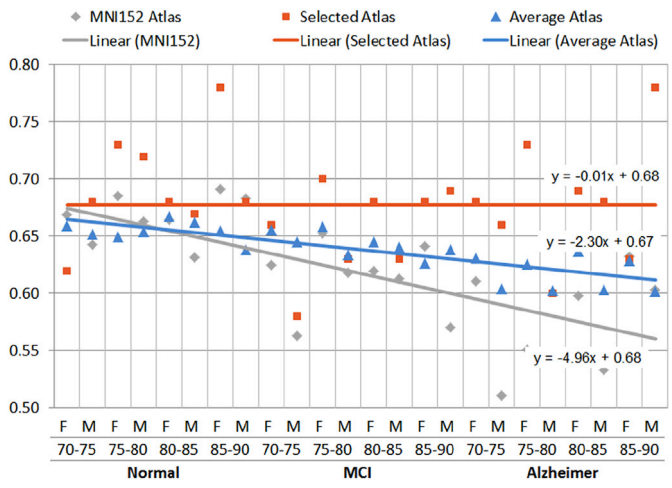

Fig. 5. Evolution of the DSC index relative to the patient's clinical situation.

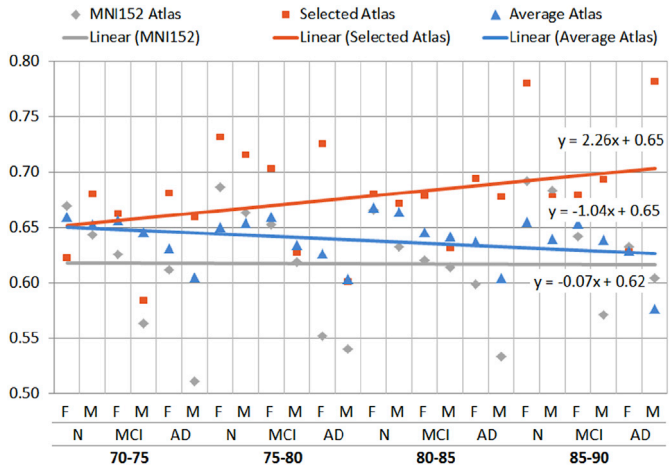

Fig. 6. Evolution of the DSC index relative to the patient's age.

3.3. Individual relevance of the selection parameters

The graphics in Fig. 7 compare the precision indexes obtained with the exclusion of each parameter to the obtained index using the three parameters together. The individuals are grouped by characteristics, the same as in Fig. 4.

When suppressing age group as a selection parameter and using only the clinical situation and gender, the average of the precision indexes fell from 0.68 ± 0.11 to 0.66 ± 0.11. This difference is significant according to a paired t-test ($\boldsymbol{p} < 0.05$, $\boldsymbol{x} = 0.68$; $\boldsymbol{\mu_0} = 0.66$; $\boldsymbol{s} = 0.11$; $\boldsymbol{n} = 1056$). The difference is more accentuated in the age group of 85–90 years. In this case, the average DSC increased from 0.66 ± 0.11 to 0.71 ± 0.10 when using age group for selection. This is an expected result, since in older individuals the hippocampal atrophy increases [50], which makes the registration process more complex, as mentioned in the previous section.

The clinical situation is also relevant for the atlas selection. Its exclusion from the selection parameters, so that only age group and gender were used, resulted in a decrease in precision from 0.68 ± 0.11 to 0.66 ± 0.11. The difference was shown to be significant through a paired t-test ($p < 0.05$; $x = 0.68$; $\mu_0 = 0.66$; $s = 0.11$; $n = 1056$). The central graphic in Fig. 7 presents the difference between including and not including this attribute as a criterion for the atlas selection.

Finally, suppressing gender as a selection parameter and using only age group and clinical situation also creates a decrease in precision from 0.68 ± 0.11 to 0.66 ± 0.10 ($p < 0.05$; $x = 0.68$; $\mu_0 = 0.66$; $s = 0.10$; $n = 1056$). However, the difference is observed only in the female group (DSC of 0.70 ± 0.10 to 0.67 ± 0.11), while in the male group there are no differences in the general average (DSC of 0.66 ± 0.11 for both).

4. Discussion

This paper evaluates the influence of using patient meta-information to better select an atlas in a hippocampus atlas-based segmentation process. By using the ADNI database composed of 353 individual test images, it was observed that an atlas selected by meta-information provides a significant increase in segmentation precision when compared to other two procedures: using a standard atlas (MNI152) or a random atlas created from a subgroup of the image set.

In this study, the method of selecting by meta-information achieved a DSC average index of 0.68, in contrast with the indexes of 0.62 obtained when using the standard MNI152 atlas and 0.64 when using an average atlas from the image set. Since our work focuses not on the segmentation method itself but on the selection technique that precedes it, it would not be valid to directly compare these indexes with state-of-the-art methods for hippocampus segmentation, like the one presented in Platero et al. [51]. Platero et al. use a multiple-atlas technique using patch-based labeling and atlas-warping that is well known to be superior to single-atlas techniques [52] such as the one applied here. However, despite the segmentation technique employed, the analysis of improvement in accuracy provided by the atlas-selection step remains valid.

Evaluating the relevance of the three parameters used in the atlas selection, namely the individual's gender, age group, and clinical situation, we can conclude that each of these characteristics is relevant as an atlas selection parameter. This occurs because the precision is negatively affected when one of the parameters is not used in the atlas selection process.

The index shows a significant increase, mainly in individuals with MCI or AD. Typically, these are the situations in which the precision indexes of segmentation methods tend to decay as the clinical situation progresses to an AD state and for which the proposed method was able to maintain a precision of 0.67. These scores are comparable to the index obtained for healthy subjects. This method is limited in the sense that atlas selection based on clinical situation cannot be used on occasions when no clinical diagnosis is available. However, when evaluating the disease progression through a series of MRI acquisitions, we can affirm that it is better to use the most recent segmentation of the patient as an atlas instead of using an average atlas.

For the evaluation of aging progression, the segmentation precision is even higher for older individuals, as can be seen from the linear regression. On average, the DCS increases from 0.66 when using an average atlas to 0.68 when using an atlas based on meta-information. Atlas distinction based on age range was also employed in a recent study on hippocampal subfield segmentation by Bender et al. [53]. Although the issue of atlas selection was not evaluated specifically, good results were obtained when segmenting using age-specific atlases for early lifespan (6–26 years) and later lifespan (62–79 years).

Atlas selection using gender information provided an increase in accuracy indexes for male individuals while keeping the precision stable for female individuals. Although this confirms the observations made by Ardekani et al. [54] when used the MIRIAD[3] database, the existence of structural sex differences in the hippocampus is a broad discussion in many other fields. Recently, Tan et al. [55] showed that this difference is not dimorphic but is only a difference in scale. This could explain the fact that gender meta-information was observed here to have the lowest relevance in atlas selection.

The non-linear registration algorithm also has a lot of influence on

[3] Minimal Interval Resonance Imaging in Alzheimer's Disease (MIRIAD). Public database provided by the Dementia Research Center of UCL (University College London).

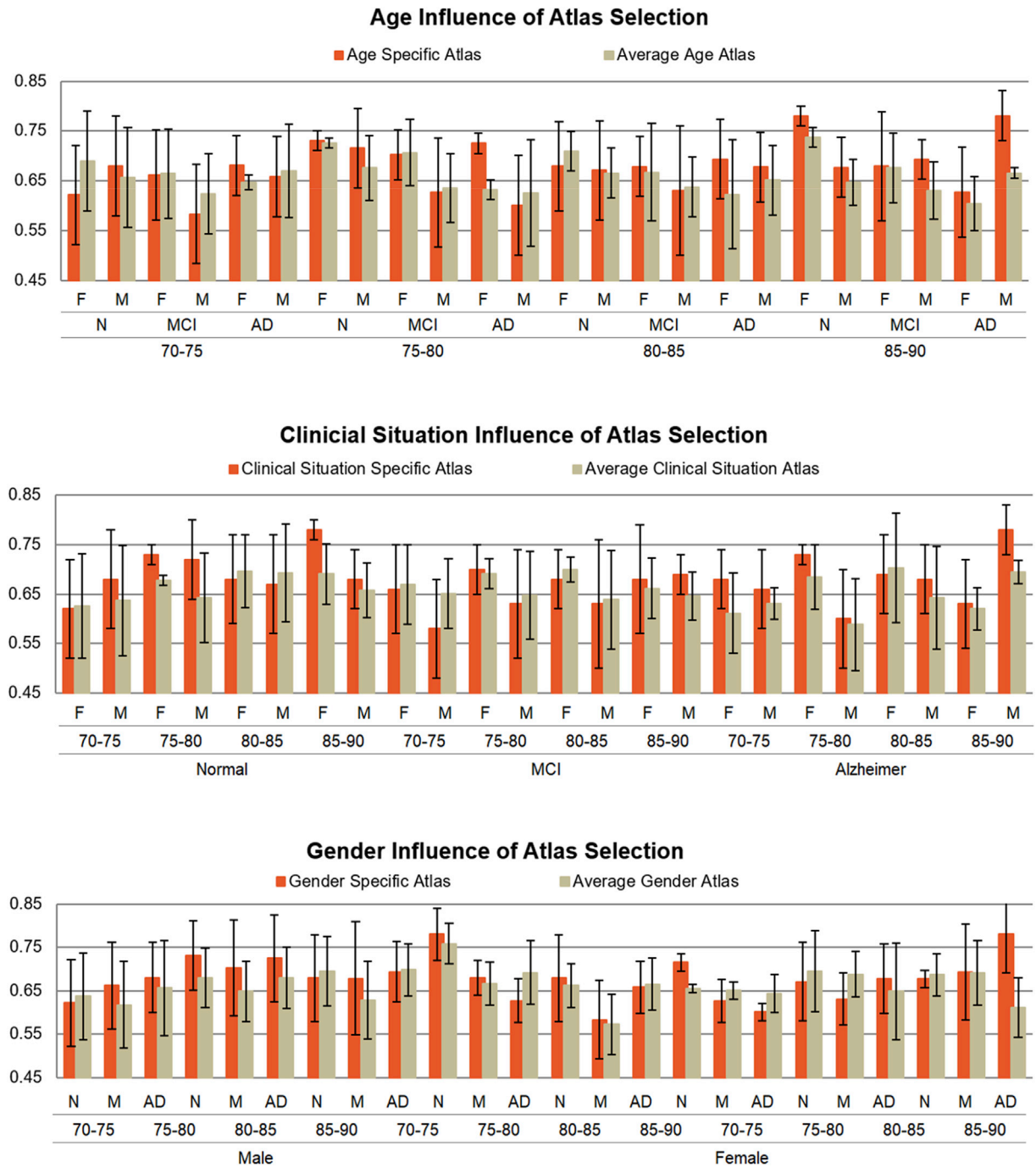

Fig. 7. Graphic comparing the influence of the selection parameters used. Error bars represent the standard deviation.

both segmentation precision and computational cost. In the implemented method, the registration precision directly influences the final segmentation precision, where an increase in the DSC index from 0.61 to 0.68 was observed when using the ART and SyN registration techniques, respectively. However, SyN does present a higher computational cost when compared to ART. This corroborates the conclusions exposed in the paper by Klein et al. [48]. More than 90% of the computation time of the implemented method is spent on the linear and non-linear registration steps. Thus, improvements in registration techniques would benefit the atlas-based segmentation methods in these processing steps.

A limitation of this study is that the "ground truth" of the hippocampus segmentation is based on segmentations performed on FreeSurfer. The ground truth image segmentation is based on a non-customized atlas. Even though all of the segmented hippocampi were inspected visually, it is possible that there are errors in some images. To our knowledge, ADNI is the largest freely available dataset with brain images of elderly subjects in which visually inspected hippocampus segmentations are included. Hence, it is currently the best dataset to be used in development of studies such as this one.

5. Conclusion

It can be concluded that the usage of atlases specific to the clinical situation, age, and gender is relevant since it improves the non-linear registration of brain images without compromising the computational cost. In general, the techniques used in the hippocampus segmentation

process steps using multiple atlases, such as label-fusion, classifiers, and optimizers, as well as the atlas selection, have reached a high level of maturity. Hence, modern methods have now achieved satisfactory precision indexes [52]. Among all the techniques used in atlas-based segmentation, the registration process is the one that burdens the method most in terms of both computational cost and segmentation precision. Therefore, future studies should continue to focus their research efforts in these areas.

Conflicts of interest

None.

Acknowledgments

Our research is partially funded by CAPES/PROSUP – Program for Supporting Gratuate Program in Private Higher Education Institutions (Grant CAPES 181/2012).

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and generous contributions from the following: AbbVie; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.compbiomed.2018.02.005.

References

[1] M.J. Cook, D.R. Fish, S.D. Shornov, K. Straughan, J.M. Stevens, Hippocampal volumetric and morphometric studies in frontal and temporal lobe epilepsy, Brain 115 (1992) 1001–1015.
[2] A. Convit, M.J.D. Leon, C. Tarshish, S.D. Santi, W. Tsui, H. Rusinek, A. George, Specific hippocampal volume reductions in individuals at risk for Alzheimer's disease, Neurobiol. Aging 18 (3) (1997) 131–138.
[3] J.G. Csernansky, S. Joshi, L. Wang, J.W. Haller, M. Gado, J.P. Miller, U. Grenander, M.I. Miller, Hippocampal morphometry in schizophrenia by high dimensional brain mapping, Proc. Natl. Acad. Sci. Unit. States Am. vol. 95 (9) (1998) 11406–11411.
[4] J.D. Bremner, M. Narayan, E.R. Anderson, L.H. Staib, H.L. Miller, D.S. Charney, Hippocampal volume reduction in major depression, Am. J. Psychiatr. 157 (1) (2000) 115–118.
[5] H.P. Blumberg, J. Kaufman, A. Martin, R. Whiteman, J.H. Zhang, J.C. Gore, D.S. Charney, J.H. Krystal, B.S. Peterson, Amygdala and hippocampal volumes in adolescents and adults with bipolar disorder, Arch. Gen. Psychiatr. 60 (12) (2003) 1201.
[6] J.D. Bremner, P. Randall, E. Vermetten, L. Staib, R.A. Bronen, C. Mazure, S. Capelli, G. McCarthy, R.B. Innis, D.S. Charney, Magnetic resonance imaging-based measurement of hippocampal volume in posttraumatic stress disorder related to childhood physical and sexual abuse—a preliminary report, Biol. Psychiatr. 41 (1) (1997) 23–32.
[7] F. Cendes, Progressive hippocampal and extrahippocampal atrophy in drug resistant epilepsy: review, Curr. Opin. Neurol. 18 (4) (2005) 173–177.
[8] B. Fischl, D.H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, A.M. Dale, Whole brain segmentation, Neuron 33 (1) (2002) 341–355.
[9] J.G. Sled, A.P. Zijdenbos, A.C. Evans, A nonparametric method for automatic correction of intensity nonuniformity in MRI data, IEEE Trans. Med. Imag. 17 (1998) 87–97.
[10] R.A. Heckemann, J.V. Hajnal, P. Aljabar, D. Rueckert, A. Hammers, Automatic anatomical brain MRI segmentation combining label propagation and decision fusion, Neuroimage 33 (10) (2006) 115–126.
[11] A. Klein, J. Andersson, B.A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G.E. Christensen, D.L. Collins, J. Gee, P. Hellier, J.H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R.P. Woods, J.J. Mann, R.V. Parsey, Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration, Neuroimage 46 (7) (2009) 786–802.
[12] M. Cabezas, A. Oliver, X. Lladó, J. Freixenet, M.B. Cuadra, A review of atlas-based segmentation for magnetic resonance brain images, Comput. Meth. Progr. Biomed. 104 (12) (2011) e158–e177.
[13] J. Barnes, J. Foster, R.G. Boyes, T. Pepple, E.K. Moore, J.M. Schott, C. Frost, R.I. Scahill, N.C. Fox, A comparison of methods for the automated calculation of volumes and atrophy rates in the hippocampus, Neuroimage 40 (5) (2008) 1655–1671.
[14] M.R. Sabuncu, B.T.T. Yeo, K.V. Leemput, B. Fischl, P. Golland, A generative model for image segmentation based on label fusion, IEEE Trans. Med. Imag. 29 (10) (2010) 1714–1729.
[15] W. Chen, S. Li, F. Jia, X. Zhang, Segmentation of hippocampus based on ROI atlas registration, in: 2011 IEEE International Symposium on it in Medicine and Education, 2011.
[16] P. Coupé, J.V. Manjón, V. Fonov, J. Pruessner, M. Robles, D.L. Collins, Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation, Neuroimage 54 (1) (2011) 940–954.
[17] J. Pipitone, M.T.M. Park, J. Winterburn, T.A. Lett, J.P. Lerch, J.C. Pruessner, M. Lepage, A.N. Voineskos, M.M. Chakravarty, Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates, Neuroimage 101 (11) (2014) 494–512.
[18] Y. Hao, T. Wang, X. Zhang, Y. Duan, C. Yu, T. Jiang, Y. Fan, Local label learning (LLL) for subcortical structure segmentation: application to hippocampus segmentation, Hum. Brain Mapp. 35 (10) (2014) 2674–2697.
[19] G. Wu, Q. Wang, D. Zhang, F. Nie, H. Huang, D. Shen, A generative probability model of joint label fusion for multi-atlas based brain segmentation, Med. Image Anal. 18 (8) (2014) 881–890.
[20] C. Platero, M.C. Tobar, J. Sanguino, O. Velasco, A new label fusion method using graph cuts: application to Hippocampus segmentation, in: IFMBE Proceedings, Springer International Publishing, 2014, pp. 174–177.
[21] Y. Song, Z. Gong, J. Yang, D. Zhao, Automatic Hippocampus segmentation of magnetic resonance imaging images using multiple atlases, Journal of Medical Imaging and Health Informatics 6 (2016) 1750–1753.
[22] L. Wang, Y. Guo, X. Cao, G. Wu, D. Shen, Consistent multi-atlas Hippocampus segmentation for longitudinal MR brain images with temporal sparse representation, in: Patch-based Techniques in Medical Imaging, Springer International Publishing, 2016, pp. 34–42.
[23] G. Wu, M. Kim, G. Sanroma, Q. Wang, B.C. Munsell, D. Shen, Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition, Neuroimage 106 (2) (2015) 34–46.
[24] J.E. Iglesias, M.R. Sabuncu, I. Aganj, P. Bhatt, C. Casillas, D. Salat, A. Boxer, B. Fischl, K.V. Leemput, An algorithm for optimal fusion of atlases with different labeling protocols, Neuroimage 106 (2) (2015) 451–463.
[25] R. Giraud, V.-T. Ta, N. Papadakis, J.V. Manjón, D.L. Collins, P. Coupé, An Optimized PatchMatch for multi-scale and multi-feature label fusion, Neuroimage 124 (1) (2016) 770–782.
[26] O.M. Benkarim, G. Piella, M.A.G. Ballester, G. Sanroma, Enhanced probabilistic label fusion by estimating label confidences through discriminative learning, in: Medical Image Computing and Computer-assisted Intervention - MICCAI 2016, Springer International Publishing, 2016, pp. 505–512.
[27] N. Bhagwat, J. Pipitone, J.L. Winterburn, T. Guo, E.G. Duerden, A.N. Voineskos, M. Lepage, S.P. Miller, J.C. Pruessner, M.M. Chakravarty, Manual-protocol inspired technique for improving automated MR image segmentation during label fusion, Front. Neurosci. 10 (7) (2016).
[28] X. Artaechevarria, A. Munoz-Barrutia, C.O. de-Solorzano, Combination strategies in multi-atlas image segmentation: application to brain MR data, IEEE Trans. Med. Imag. 28 (8) (2009) 1266–1277.
[29] P. Aljabar, R.A. Heckemann, A. Hammers, J.V. Hajnal, D. Rueckert, Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy, Neuroimage 46 (7) (2009) 726–738.
[30] K.K. Leung, J. Barnes, G.R. Ridgway, J.W. Bartlett, M.J. Clarkson, K. Macdonald, N. Schuff, N.C. Fox, S. Ourselin, Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease, Neuroimage 51 (7) (2010) 1345–1359.
[31] J.M.P. Lötjönen, R. Wolz, J.R. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, D. Rueckert, Fast and robust multi-atlas segmentation of brain magnetic resonance images, Neuroimage 49 (2010) 2352–2365.
[32] A.R. Khan, N. Cherbuin, W. Wen, K.J. Anstey, P. Sachdev, M.F. Beg, Optimal weights for local multi-atlas fusion using supervised learning and dynamic information (SuperDyn): validation on hippocampus segmentation, Neuroimage 56 (5) (2011) 126–139.

[33] J. Lötjönen, R. Wolz, J. Koikkalainen, V. Julkunen, L. Thurfjell, R. Lundqvist, G. Waldemar, H. Soininen, D. Rueckert, Fast and robust extraction of hippocampus from MR images for diagnostics of Alzheimer's disease, Neuroimage 56 (5) (2011) 185–196.
[34] A. Akhondi-Asl, K. Jafari-Khouzani, K. Elisevich, H. Soltanian-Zadeh, Hippocampal volumetry for lateralization of temporal lobe epilepsy: automated versus manual methods, Neuroimage 54 (1) (2011) S218–S226.
[35] M.J. Cardoso, K. Leung, M. Modat, S. Keihaninejad, D. Cash, J. Barnes, N.C. Fox, S. Ourselin, STEPS: similarity and Truth Estimation for Propagated Segmentations and its application to hippocampal segmentation and brain parcelation, Med. Image Anal. 17 (8) (2013) 671–684.
[36] J. Ma, H.T. Ma, H. Li, C. Ye, D. Wu, X. Tang, M. Miller, S. Mori, A fast atlas pre-selection procedure for multi-atlas based brain segmentation, in: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015.
[37] A.S. Korsager, V. Fortunati, F. Lijn, J. Carl, W. Niessen, L.R. {\O}stergaard, T. Walsum, The use of atlas registration and graph cuts for prostate segmentation in magnetic resonance images, Med. Phys. 42 (3) (2015) 1614–1624.
[38] J.V. Velde, J. Wouters, T. Vercauteren, W.D. Gersem, E. Achten, W.D. Neve, T.V. Hoof, The effect of morphometric atlas selection on multi-atlas-based automatic brachial plexus segmentation, Radiat. Oncol. 10 (12) (2015) 1.
[39] L.R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (7) (1945) 297–302.
[40] M. Jenkinson, S. Smith, A global optimisation method for robust affine registration of brain images, Med. Image Anal. 5 (6) (2001) 143–156.
[41] V. Fonov, A.C. Evans, K. Botteron, C.R. Almli, R.C. McKinstry, D.L. Collins, Unbiased average age-appropriate atlases for pediatric studies, Neuroimage 54 (1) (2011) 313–327.
[42] S.M. Smith, Fast robust automated brain extraction, Hum. Brain Mapp. 17 (11) (2002) 143–155.
[43] M. Jenkinson, M. Pechaud, S. Smith, BET2: MR-based estimation of brain, skull and scalp surfaces, in: Eleventh Annual Meeting of the Organization for Human Brain Mapping, 2005.
[44] M. Jenkinson, C.F. Beckmann, T.E.J. Behrens, M.W. Woolrich, S.M. Smith, FSL, Neuroimage 62 (8) (2012) 782–790.
[45] B.A. Ardekani, M. Braun, B.F. Hutton, I. Kanno, H. Iida, A fully automatic multimodality image registration algorithm, J. Comput. Assist. Tomogr. 19 (7) (1995) 615–623.
[46] B.A. Ardekani, S. Guckemus, A. Bachman, M.J. Hoptman, M. Wojtaszek, J. Nierenberg, Quantitative comparison of algorithms for inter-subject registration of 3D volumetric brain MRI scans, J. Neurosci. Meth. 142 (3) (2005) 67–76.
[47] B.B. Avants, C.L. Epstein, M. Grossman, J.C. Gee, Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain, Med. Image Anal. 12 (2) (2008) 26–41.
[48] A. Klein, S.S. Ghosh, B. Avants, B.T.T. Yeo, B. Fischl, B. Ardekani, J.C. Gee, J.J. Mann, R.V. Parsey, Evaluation of volume-based and surface-based brain image registration methods, Neuroimage 51 (5) (2010) 214–220.
[49] G.B. Frisoni, M.P. Laakso, A. Beltramello, C. Geroldi, A. Bianchetti, H. Soininen, M. Trabucchi, Hippocampal and entorhinal cortex atrophy in frontotemporal dementia and Alzheimer's disease, Neurology 52 (1) (1999), 91–91.
[50] J.C. Smith, K.A. Nielson, J.L. Woodard, M. Seidenberg, S. Durgerian, K.E. Hazlett, C.M. Figueroa, C.C. Kandah, C.D. Kay, M.A. Matthews, S.M. Rao, Physical activity reduces hippocampal atrophy in elders at genetic risk for Alzheimer's disease, Front. Aging Neurosci. 6 (4) (2014).
[51] C. Platero, M.C. Tobar, A fast approach for hippocampal segmentation from T1-MRI for predicting progression in Alzheimer's disease from elderly controls, J. Neurosci. Meth. 270 (2016) 61–75.
[52] V. Dill, A.R. Franco, M.S. Pinho, Automated methods for Hippocampus segmentation: the evolution and a review of the state of the art, Neuroinformatics 13 (10) (2015) 133–150.
[53] A.R. Bender, A. Keresztes, N.C. Bodammer, Y.L. Shing, M. Werkle-Bergner, A.M. Daugherty, Q. Yu, S. Kuhn, U. Lindenberger, N. Raz, Optimization and validation of automated hippocampal subfield segmentation across the lifespan, Hum. Brain Mapp. 39 (2) (2018) 916–931.
[54] B.A. Ardekani, A. Convit, A.H. Bachman, Analysis of the MIRIAD data shows sex differences in hippocampal atrophy progression, J. Alzheim. Dis. 50 (2) (2016) 847–857.
[55] A. Tan, W. Ma, A. Vira, D. Marwha, L. Eliot, The human hippocampus is not sexually-dimorphic: meta-analysis of structural MRI volumes, Neuroimage 124 (2016) 350–366.