

NABS: Non-local Automatic Brain Hemisphere Segmentation

José E. Romero¹, José V. Manjón¹, Jussi Tohka², Pierrick Coupé³, Montserrat Robles¹

¹ Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA),
Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain.

² Department of Signal Processing, Tampere University of Technology, P.O. Box 553, 33101, Tampere,
Finland.

³ Laboratoire Bordelais de Recherche en Informatique, Unité Mixte de Recherche CNRS (UMR 5800),
PICTURA Research Group, 351, cours de la Libération F-33405 Talence cedex, France.

Abstract

In this paper, we propose an automatic method to segment the five main brain sub-regions (i.e. left/right hemispheres, left/right cerebellum and brainstem) from magnetic resonance images. The proposed method uses a library of pre-labeled brain images in a stereotactic space in combination with a non local label fusion scheme for segmentation. The main novelty of the proposed method is the use of a multi-label block-wise label fusion strategy specifically designed to deal with the classification of main brain sub-volumes that process only specific parts of the brain images significantly reducing the computational burden. The proposed method has been quantitatively evaluated against manual segmentations. The evaluation showed that the proposed method was faster while producing more accurate segmentations than a current state-of-the-art method. We also present evidences suggesting that the proposed method was more robust against brain pathologies than the compared method. Finally, we demonstrate the clinical value of our method compared to the state-of-the-art approach in terms of the asymmetry quantification in Alzheimer's disease.

Keywords: Brain segmentation, asymmetry, brain volume analysis, patch-based segmentation, MRI.

1. Introduction

The hemisphere segmentation of magnetic resonance (MR) brain images refers to the extraction of the left and right cerebral and possibly cerebellar hemispheres from images. It has been applied to assess brain asymmetries [1] to elucidate functional brain lateralization [2], natural brain development and neurodegeneration [3], and the effects of diseases on the human brain. Particularly, many brain diseases such as Alzheimer's disease [4], Autism [5] and Schizophrenia [6, 7] have been associated with abnormal asymmetry patterns. Furthermore, hemisphere segmentation is required in order to view the medial surface of the cerebral hemispheres, and many important brain structures such as the medial temporal lobe, cingulum, and large portions of the occipital, parietal and frontal lobes can only be visualized on the medial hemispheric surface [8]. In addition, hemispheric shape analysis has been used to compare populations according to different criteria [9, 10].

Several automatic strategies have been developed for hemisphere segmentation and compartmental segmentation, the latter referring to the division of the brain into the left and right cerebral hemispheres (CH), Cerebellum (CB), and Brainstem (BS). As demonstrated by Zhao et al. [11] hemisphere segmentation approaches based on mid-sagittal plane extraction or linear registration [12, 13] tend lead to inaccurate segmentation results because the brain is not perfectly symmetric. The same holds true for approaches used in FreeSurfer [14] and BrainVoyager [15], which apply two cutting planes, one separating the left and right hemispheres and the other separating the cerebrum from the CB and BS. The methods aiming for more accurate hemisphere/compartmental segmentations can be based on the search of the segmentation surface between the hemispheres, usually by optimizing some cost function [8, 16], nonlinear registration [17, 18] or structure-reconstruction where one first finds seed voxels representing the hemispheres (and cerebellum) and then reconstructs the hemispheres from these seed voxels [11, 19, 20].

The use of pre-segmented templates [17] has been demonstrated to be a powerful tool to automatically delineate brain compartments. Such methods use prior knowledge about the human brain anatomy in the form of a pre-labeled template segmentation. However, since these approaches use nonlinear registration with a single template, they are limited by the problem of trying to match dissimilar anatomical patterns resulting from inter-subject variability. To alleviate

this problem, label fusion techniques based on multiple templates have been proposed. These techniques take advantage of the inter-subject variability of the pre-selected templates and combine its propagated labels in order to decide the final label for every voxel of the target image. Earliest methods like majority voting [A, B, C] consists of calculating the most frequent label from the templates for a given voxel ignoring the image intensities. Recent methods [21, 22] can produce high quality segmentations even when only linear transformations are used as the method described in [23] which is not limited to assign a single label from every template. For every target voxel it performs a patch matching along a search area surrounding the voxel selecting several label candidates from every template. This matching is based on intensity and contrast to ensure a good contribution. Then the final label is selected by a weighted voting. It is worth noting that label fusion techniques have been used in combination with non-linear registrations providing state of the art results at the expense of a higher computational load [24].

In this paper, a new cost-efficient segmentation method to accurately separate left and right brain hemispheres, cerebellum and brainstem is proposed. As originally proposed by Coupé et al. [23], our method is based on a nonlocal means label fusion where labels from multiple templates are weighted according to the Euclidean distance between patch intensities. Using this approach we avoid the one-to-one matching assumption of nonlinear registration label fusion methods by enabling a one-to-many matching, therefore reducing the segmentation errors [23, 25, 26] by better managing the inherent inter-subject variability of human brain anatomy.

The main novelties of the proposed method are the use of an optimized multi-label block-wise label fusion strategy specially designed to deal with the classification of main brain compartments which significantly reduces the method complexity and an extensive validation of the proposed methodology.

2. Materials and methods

The method proposed in this paper uses a library of manually labeled brain templates as a prior knowledge base to drive the segmentation process. Such library has been constructed using a publicly available IXI dataset [<http://www.brain-development.org/>]. This dataset contains images from nearly 600 healthy subjects that have been collected at three different hospitals of London. In this section, the template library construction is first detailed and then the proposed method is presented.

2.1. Template library construction

To generate the template library, 30 T1-weighted images acquired on 1.5 T and 3 T machines from different vendors were randomly selected from the IXI dataset [<http://www.brain-development.org/>] (age range: 24-75 years, volume size: 256 x 125 x 150 voxels and 1x1x1.2 mm³ voxel resolution). The images were preprocessed as summarized in Figure 1.

2.1.1. Denoising

All images in the database were first denoised using the Spatially Adaptive Non-Local Means (SANLM) Filter [27] to enhance the image quality. The SANLM filter can deal with noise levels spatially varying across the image. The noise level is locally estimated which makes it ideal to process data with either stationary or spatially varying noise fields in a fully automatic manner.

2.1.2. MNI space registration

For the segmentation process, the templates and the image to be segmented have to be placed in the same stereotactic space. Therefore, a spatial normalization based on a linear registration to the Montreal Neurological Institute (MNI152) space was performed using SPM8 [28] routines. After the registration process, all the template images had a common size of 181x217x181 voxels with 1 mm³ voxel resolution.

2.1.3. Brain extraction and Inhomogeneity correction

Since we are only interested on intracranial tissues, a brain extraction operation was performed to exclude the extracranial tissues from the images. For this purpose, the VBM8 software was used [<http://dbm.neuro.uni-jena.de/>]. We used VBM8 toolbox since it is a fast, robust and widely used package for MRI segmentation. In addition, the inhomogeneity of the images were corrected using VBM8.

2.1.4. Intensity normalization

As the proposed method is based on the estimation of image similarities using intensity-based measures, every image in the library must be intensity normalized. More specifically, the intracranial tissue types (White Matter(WM), Grey Matter(GM) and Cerebrospinal Fluid(CSF)) must have consistent mean intensities across the images). For this purpose, mean intensity values of the intracranial tissue types were estimated using the TMS method [29]. Finally, a piecewise linear intensity mapping [30] was performed in such a way that WM had an average intensity of 250, GM 150 and CSF of 50 for all subjects in the library.

2.1.5. Manual labeling

Manual labeling of the 5 different considered brain compartments (Left Hemisphere, Right Hemisphere, Left Cerebellum, Right Cerebellum and Brainstem) was performed by an expert in brain anatomy using ITK-SNAP software [31]. To speed up the manual labeling process we first obtained an initial labeling by applying the ADisc method [11] and then we manually corrected the segmentation errors.

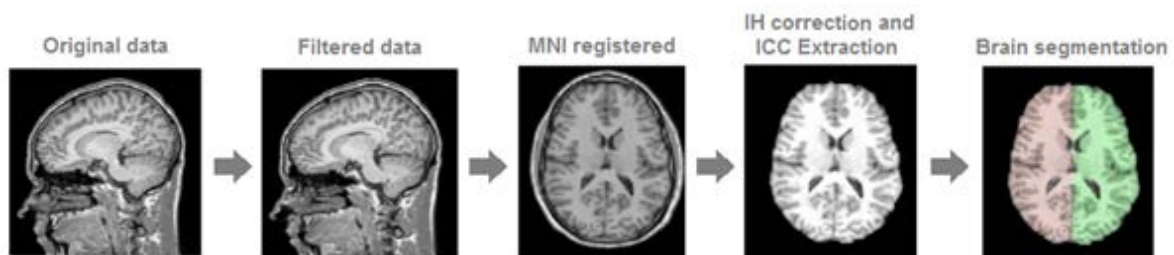


Figure 1. The template library construction pipeline.

2.2. Proposed Method

The proposed method is based on the Nonlocal means label fusion originally proposed by Coupé et al. [23]. However, some modifications have been applied in order to adapt it to nature of the hemisphere/compartamental segmentation problem. We will refer to our proposed method as NABS (**N**on-local **A**utomatic **B**rain hemisphere **S**egmentation). To segment the different brain regions, NABS first preprocesses the image data using the pipeline described in the previous section and then our label fusion scheme to label each brain voxel is applied. Below, we will first describe the main features of the non-local label fusion method [23] and then we will highlight the modifications proposed in this paper to solve the hemispheric brain segmentation problem.

2.2.1. Nonlocal means estimator for binary segmentation

For each voxel x_i from the image to be segmented the method estimates the final label $L(x_i)$ by performing a weighted label fusion $v(x_i)$ of all surrounding samples inside the search volume V_i from the N subjects of the library:

$$v(x_i) = \frac{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j}) y_{s,j}}{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j})} \quad (1)$$

where $y_{s,j}$ is a possible label from the voxel $x_{s,j}$ at the position j in the subject s and $w(x_i, x_{s,j})$ is the weight calculated by patch comparison which is computed depending on the similarity of the surrounding patch for x_i and for $x_{s,j}$:

$$w(x_i, x_{s,j}) = \begin{cases} \exp \frac{-\|P(x_i) - P(x_{s,j})\|_2^2}{h^2}, & ss < th \\ 0, & otherwise \end{cases} \quad (2)$$

where $P(x_i)$ is the patch around the voxel x_i , $P(x_{s,j})$ is the patch around the voxel x_j in the templates and $\|.\|_2$ is the normalized L2 norm (normalized by the number of elements) calculated from the distance between each pair of voxels from both patches $P(x_i)$ and $P(x_{s,j})$. The structural similarity index [32] ss is calculated as :

$$ss = \frac{2\mu_i\mu_{s,j}}{\mu_i^2 + \mu_{s,j}^2} \cdot \frac{2\sigma_i\sigma_{s,j}}{\sigma_i^2 + \sigma_{s,j}^2} \quad (3)$$

where μ_i and σ_i are the mean and standard deviation of the patch $P(x_i)$ surrounding x_i . If ss (is less than a threshold th (we set $th = 0.97$ in this work) then w is not computed in order to save useless computation time. The normalization parameter h was estimated as in Coupé et al. [23] based on the minimal patch distance in the search area.

Finally, the final label $L(x_i)$ is computed as:

$$L(x_i) = \begin{cases} 1 & v(x_i) \geq 0.5 \\ 0 & v(x_i) < 0.5 \end{cases} \quad (4)$$

2.2.2. Block-wise Nonlocal means estimator multi-label segmentation

In order to segment brain hemispheres, cerebellum and brainstem, we must modify equation (1) since it is designed for binary labels $\{0,1\}$ and the proposed method requires multiple labels. In this particular case the set of labels is $\{0,1,2,3,4,5\}$ and Equation (1) is modified to:

$$v(x_i, k) = \frac{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j}) \delta(k, y_{s,j})}{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j})} \quad (5)$$

where δ is the Kronecker's delta function and $k = 0,1,\dots,5$. Finally, for each voxel, the most voted label is assigned:

$$L(x_i) = \arg \max_k v(x_i, k) \quad (6)$$

We can reduce the computational burden of the method by using a block-based approach rather than a voxel-wise one as previously proposed for the MRI denoising [33]. This means that we label the whole block of voxels simultaneously instead of performing this operation in a voxel by voxel basis. Here, we consider blocks of $5 \times 5 \times 5$ voxels centered on every other voxel in each direction.

Recently, Rousseau et al. [26] applied the block-wise strategy proposed by Coupe et al. [33] to nonlocal means label fusion resulting in slightly better segmentations than those obtained by a voxel-based implementation. In fact, the block-based approach not only provides a reduction of the computational load by processing only every other in all three dimensions as done by Coupe et al. [33] but it also introduces an implicit regularization in the labels due to the overlap between blocks which increases the number of patches involved in the voting process for each voxel.

Therefore, the new vote definition can be defined as:

$$v(B(x_i), k) = \frac{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j}) \delta(k, B(y_{s,j}))}{\sum_{s=1}^N \sum_{j \in V_i} w(x_i, x_{s,j})} \quad (7)$$

where $B(x_i)$ is a 5x5x5 voxel block which is labeled at the same time. In above, we have slightly abused notation: $\delta(k, B(y_{s,j}))$ returns a 5x5x5 voxel block, where a label block $B(y_{s,j})$ is compared to the label k treating each voxel separately and $v(B(x_i), k)$ is a 5x5x5 vote block centered at the voxel x_i . Finally, the vote count $v(x_i, k)$ for the voxel x_i is obtained by summing votes in every block containing x_i , i.e.

$$v(x_i, k) = \sum_{x_i \in B(z)} [v(B(z), k)]_i \quad (8)$$

where $[v(B(z), k)]_i$ refers to the element corresponding to x_i in the block $v(B(z), k)$ and the label $L(x_i)$ is decided as in Eq. (6).

2.2.3. Processing area reduction and location based labeling

Even with the block-wise approach, the computational burden of the method is still very high since the method has to label all brain voxels. However, we can take advantage from the fact that many of the required voxel-labels can be estimated based on the voxel position in the stereotactic MNI space. This means that voxels located far from the boundaries between different brain regions can be directly labeled by the most frequent label at that position within the library of templates. Only voxels around the boundaries need to be visited to decide the proper label.

To this end, a pre-calculated mask is obtained using the 30 templates from the template library where voxels near the boundaries between the brain regions are marked as *uncertain* and will be

visited during the segmentation process. The remaining *certain* voxels are labeled at each position with the mode of the label distribution from all the cases in the library. We obtain the uncertain region by dilating each label separately (using a 3D spherical kernel of 5 voxels radius for further security) and fusing all resulting intersections to create the uncertainty mask (see figure 2).

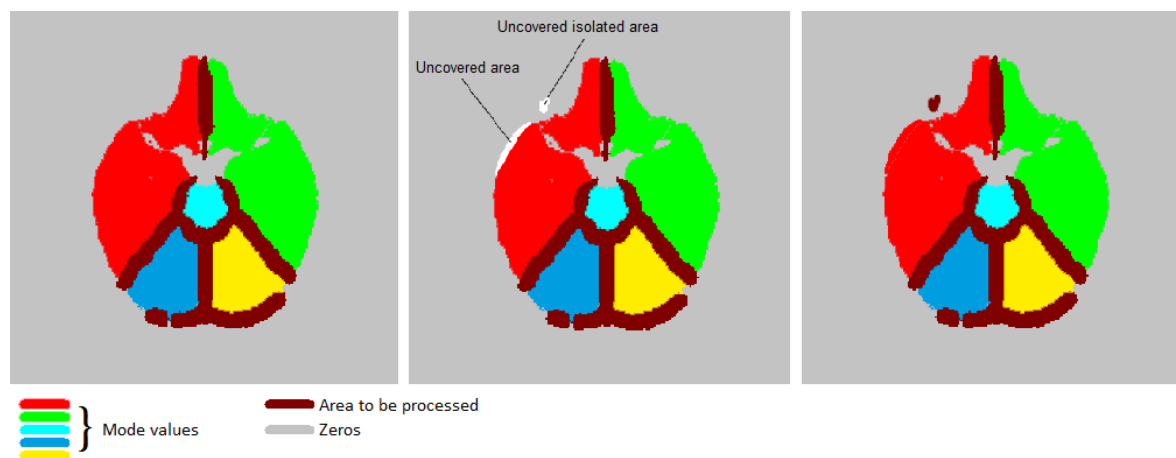


Figure 2. Left: The pre-calculated mask from the library. Center: The mask over a new case (notice that some areas remain uncovered). Right: The final adjusted mask after label dilation. Mode values stands for mode value of the labels for each voxel across all cases of the library.

Since this pre-calculated mask provides a default labeling for a new case to be segmented at the *certain* areas (grey and white matter voxels not belonging to uncertain areas are labeled with the tag associated for that position at the pre-calculated mask) it is required to properly adapt the default mask to the new case. To achieve this, after the mode based labeling, connected areas not belonging to the uncertain area are filled by using a region growing process where every non assigned voxel is labeled using the label of the nearest labeled area. Finally, since some isolated voxels may remain uncovered after the region growing process they are marked as uncertain so they will be visited during the segmentation (see figure 2 for details).

2.2.4. Pre-selection

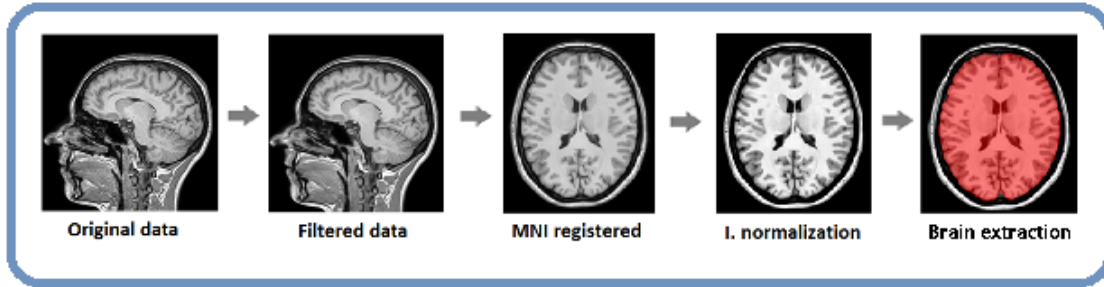
As shown in previous works [22, 23], the pre-selection of the most similar templates within the library to the target brain can be performed before segmentation to select only the most adequate samples from the library and also to reduce the computational load of the method. To measure the similarity between the case to be segmented and each one of the templates in the library the mean absolute difference (MAD) was used:

$$MAD(s) = \frac{1}{|\Omega|} \sum_{i=1}^{|\Omega|} |x_{s,i} - x_i| \quad (9)$$

where $|\Omega|$ is the total number of voxels.

The whole scheme of NABS method including the preprocessing and labeling steps are outlined in Figure 3.

Preprocessing pipeline



Labelling

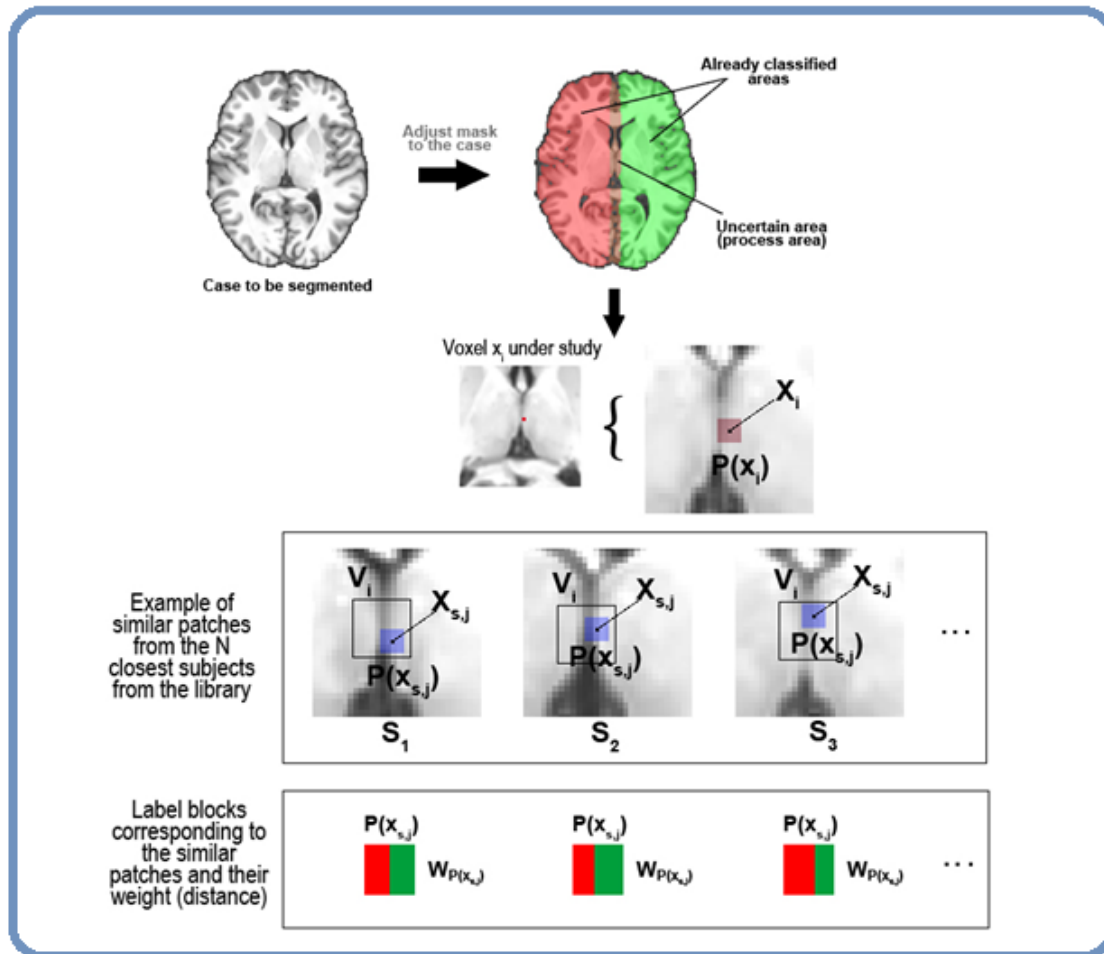


Figure 3. Overview of the preprocessing pipeline and the labelling method showing the steps involved in the selection of label blocks $P(x_{s,j})$ from the search area volume V_i along the N nearest cases from the library. The weight $W_{P(x_{s,j})}$ stores the value calculated from the distance between $P(x_{s,j})$ and $P(x_i)$ which will be used as a vote for every voxel.

3. Experiments and results

To validate the proposed method, we performed a leave-one-out cross validation using the 30 case examples of the library. Cohen's Kappa metric [34] was used to measure the similarity between the automatic segmentation and the corresponding manual one. Kappa metric was measured only with the brain mask area. All the experiments were performed using MATLAB 2009b (Mathworks Inc.).

3.1. Parameter selection

An exhaustive search of the best parameter values for the method was performed by measuring the quality of the segmentation as a function of the most important parameters (search volume size and number of selected templates).

The proposed method looks for similar patterns within a search volume to drive the segmentation process. The size of this search volume has an important effect on the quality of the results. Thus, we tested different search volume sizes to find the optimum size in terms of accuracy and computational time. Figure 4 shows average Cohen's kappa values across 30 images of the library as a function of the radius of search volume (search volume is defined as a $2 \times \text{radius} + 1$ voxels side cube). In this experiment, all library templates were used in the leave-one-out setting (the number of templates was 29).

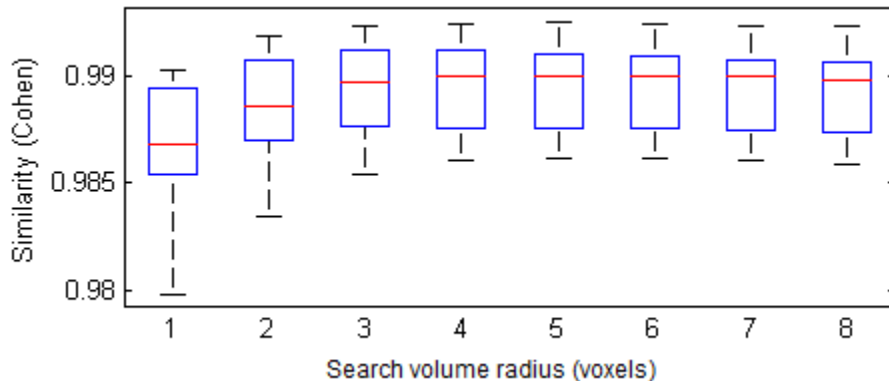


Figure 4. Impact of the search volume radius on the similarity.

Similarly to the Coupe findings [23], our results confirmed that a radius of 4 voxels is a good compromise between quality and computational load. Therefore, we use set 4 voxels as the default value of the search volume radius.

We performed another experiment to find the optimum number of pre-selected cases for balancing the segmentation accuracy and the computational load. We estimated the optimal number of templates from the library by measuring the Cohen’s Kappa index as a function of the number of selected cases from the library. Figure 5 shows Box plots of the Cohen’s Kappa index as a function of the number of cases used to perform the segmentation. As can be noticed, the proposed method reached a stable plateau around N=20 cases which is also in good agreement with previously published results [21, 23]. Thus, we set N=20 as the default parameter in NABS method.

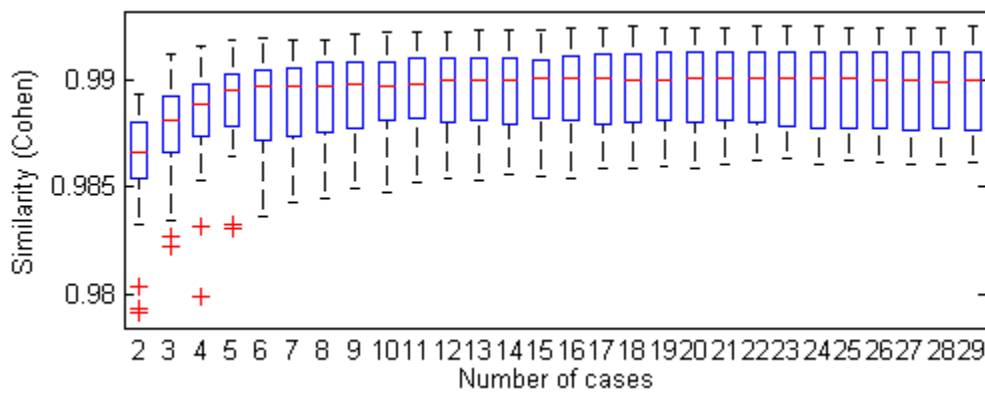


Figure 5. Impact of the number of templates on the segmentation accuracy of NABS. Each point represents the distribution of Cohen's Kappa index of the 30 library cases as a function of the number of most similar samples used for its segmentation.

3.2. Block-wise implementation evaluation

As explained in section 2.2.2, a block-wise approach has been used in the Non-local means label fusion method to improve the segmentation accuracy and drastically reduce the computational load. We compared voxel-wise and block-wise versions to find out the differences of the two versions. The mean accuracy obtained with the voxel-wise version was 0.9889 ± 0.0023 while the block-wise obtained 0.9895 ± 0.0020 (see figure 6). A paired t-test was performed to probe the

statistical significance of the differences with a p-value=0.0019. Also a significant improvement on the average computation time was obtained (185 seconds for the voxel-wise implementation compared to 60 seconds for the block-wise implementation).

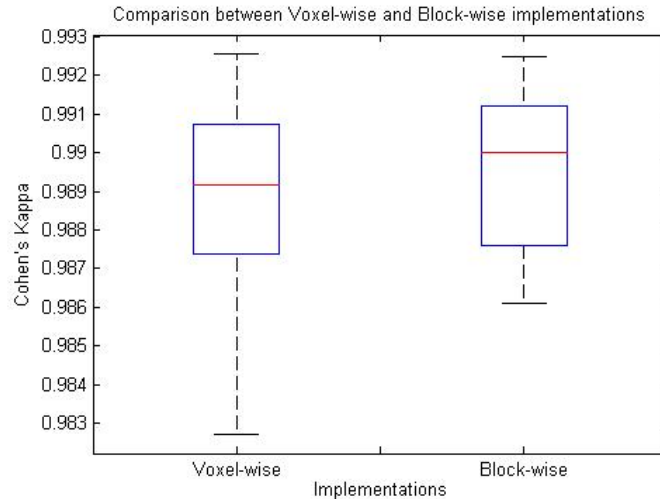


Figure 6. A comparison between voxel-wise and block-wise implementations in term of segmentation accuracy.

3.3. Comparison with Adaptive Disconnection method

In order to evaluate our proposed method we compared it with the Adaptive Disconnection (ADisc) method [11] that represents the *state-of-the-art* on hemisphere segmentation. ADisc method has been shown to outperform brainVISA [19], linear registration based mid-sagittal plane approach and non-linear registration based segmentation. The same image preprocessing steps were applied for the two methods compared to avoid any bias in the method comparison. The Matlab MEX source code of Adaptive Disconnection method was obtained from <http://www.cs.tut.fi/~jupeto/software.html>).

The ADisc method is based on the iterated use of the shape bottlenecks algorithm [35] which, by modeling the information transfer in a binary image using Laplace's equation, can detect the bridge-like connections between two parts of a complex object. First, by using the shape bottlenecks algorithm, ADisc detects the connections between the CH, CB, and BS in the WM mask of the brain. Then, it reconstructs the three compartments based on a specific region growing

algorithm. As the final step, CB and CH are divided into left and right hemispheres, again by using the shape bottlenecks algorithm.

3.3.1. Segmentation accuracy

The 30 library cases were processed using both NABS and ADisc methods. The global mean Cohen's kappa similarity value (measured using all brain voxels) obtained by ADisc was 0.9868 ± 0.0017 while the mean similarity value obtained by NABS was 0.9962 ± 0.0014 . Focusing only on the ROI corresponding to the *visited* areas (see figure 2), the mean similarity value obtained by ADisc was 0.8912 ± 0.0137 and the mean value obtained by NABS was 0.9729 ± 0.0103 . As can be noticed in figure 7 the differences between the methods were significant ($p = 3.7 \times 10^{-20}$ for global results and $p = 1.6 \times 10^{-22}$ for ROI results using a paired t-test).

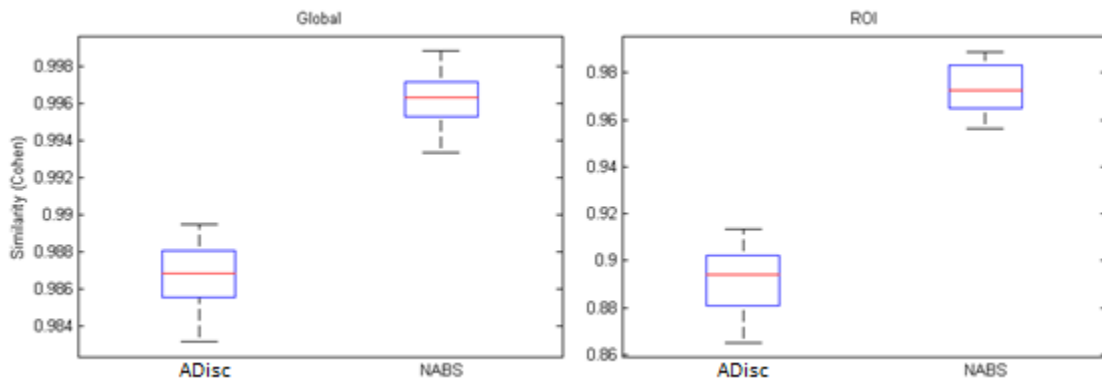


Figure 7. Box Whisker diagram comparing Adaptive disconnection (ADisc) and our method (NABS).

To perform a fair method comparison, only voxels labeled as brain (GM+WM) in both methods were taken into account to measure kappa coefficients (note the differences in NABS method in figures 6 and 7). This was necessary since the methods had slightly different thresholds for brain/non-brain intensities (this can be noticed in figure 8). A visual comparison of the segmentations by the different methods is provided also in figure 8 to better illustrate the differences between the methods compared.

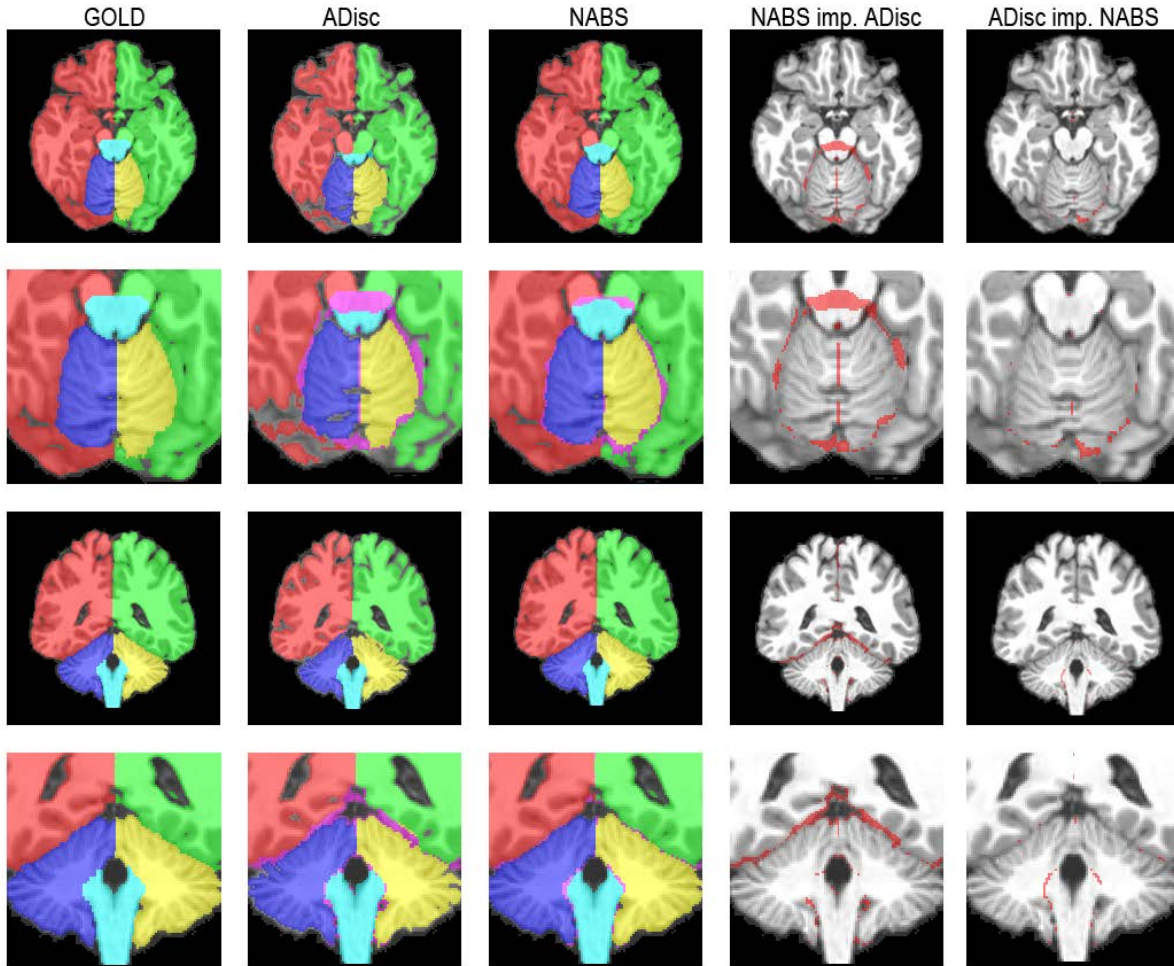


Figure 8. Comparison between methods. Purple areas in columns 2 and 3 show ADisc and NABS errors. Highlighted regions in columns 4 and 5 show areas where NABS improves ADisc and vice versa. GOLD stands for the gold standard.

To better understand where the differences between segmentations by ADisc and NABS were localized, we measured segmentation accuracy also for every sub-region separately. As previously, we measured the segmentation accuracy both globally and locally (visited areas ROI). In this case, we used Dice's coefficient [36] instead of Cohen's kappa because binary labels were compared. These results are provided in Table 1.

Table 1. Mean Dice's coefficient results for different sub-regions.

	Method	Left CH	Right CH	Left CB	Right CB	BS
GLOBAL	ADisc	0.9961 ±0.0007	0.9959 ±0.0006	0.9717 ±0.0074	0.9705 ±0.0068	0.9315 ±0.0213
	NABS	0.9985 ±0.0006	0.9984 ±0.0006	0.9906 ±0.0048	0.9900 ±0.0053	0.9817 ±0.0074
ROI	ADisc	0.9467 ±0.0092	0.9490 ±0.0112	0.8698 ±0.0476	0.8698 ±0.0444	0.8189 ±0.0691
	NABS	0.9818 ±0.0042	0.9817 ±0.0057	0.9648 ±0.0157	0.9617 ±0.016	0.9445 ±0.0267

As can be noticed in Table 1, CB and BS presented a bigger improvement for both global and ROI measurements (all differences were found to be statistically significant with a p-value<0.05). This clearly explains that main benefits of the proposed method can be found at those areas where NABS was able to better deal with partial volume affected areas such as cerebellum boundaries (see also the improved areas in figure 8).

3.3.2. Asymmetry estimation accuracy

As one of the main applications for hemisphere segmentation is brain asymmetries estimation, we studied the impact of the different methods on this measure. We estimated the asymmetry index as

$$Asymmetry = \frac{R - L}{R + L} \quad (10)$$

where R and L are the volumes corresponding to the right and left halves of the structure of interest.

Both cerebrum and cerebellum asymmetry indexes were obtained from the segmentation results of the compared automated methods (NABS and Adaptive disconnection) and compared to asymmetry indexes computed based on the manual segmentations. The results of asymmetry estimation are shown in figure 9.

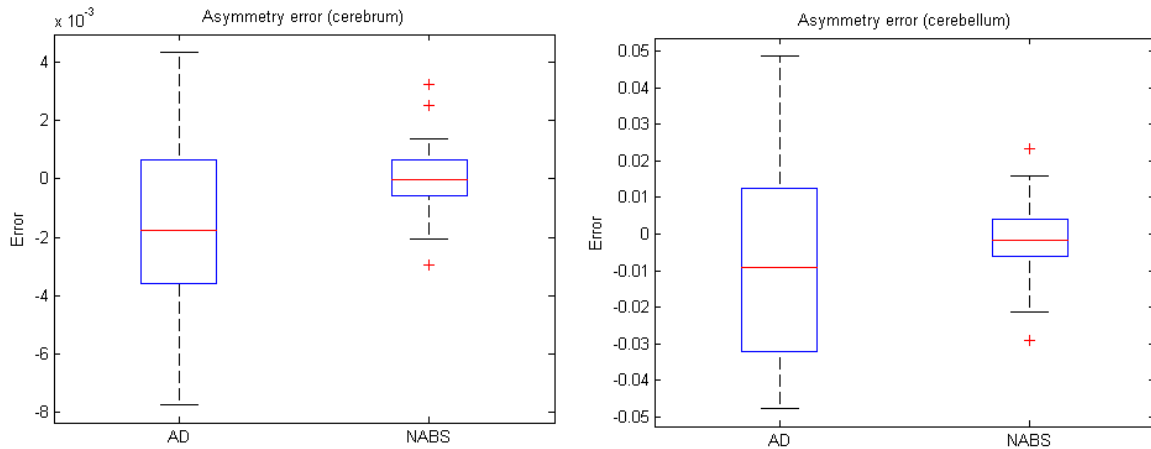


Figure 9. Asymmetry error for cerebrum and cerebellum using ADisc and NABS.

As can be noticed in figure 9, ADisc method slightly under-estimated the asymmetry for both sub-regions (cerebrum and cerebellum) compared to the manual segmentation while the average error of NABS was closer to zero. Furthermore, the variance of errors was lower with NABS than with ADisc. The mean error for ADisc was -0.0016 ± 0.0029 and for NABS it was 0.00004 ± 0.0012 . The differences between methods were found to be statistically significant (paired-t-test $p < 0.05$).

3.3.3. Simulated lesion test

In order to estimate the robustness of the compared methods when facing abnormal cases, we introduced a simulated white matter lesion in a right hemisphere of an image (excluded from the library) by reducing the mean intensity within the lesion to 50% of the original mean intensity (from 250 to 125). The simulated lesion had the volume of 12 cm^3 . Figure 10 shows the results of the different methods for this simulated case. As it can be observed, NABS segmentation was not affected by the presence of the lesion. Instead, ADisc failed the core of the lesion and, even more interestingly, introduced an alteration of the boundary between left and right hemispheres (probably due to the alteration of the potential map of ADisc method). The similarity value obtained for this case was 0.8961 for ADisc and 0.9656 for NABS (ROI).

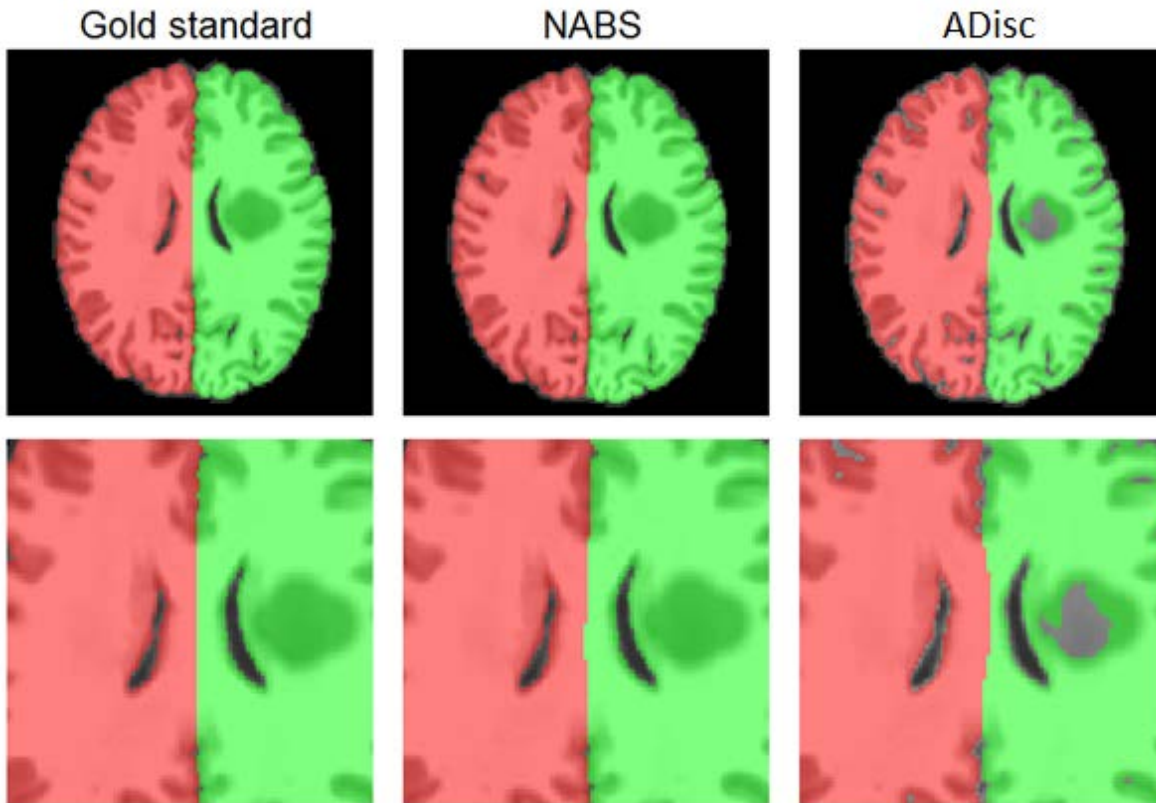


Figure 10. From left to right: Gold standard, NABS and ADisc example slice of segmentation results and their corresponding close-ups.

3.3.4. Asymmetry comparison between normal aging and mild dementia groups

It has been shown that patients with AD present a higher level of brain asymmetry than normal, age-matched controls [Thompson et al., 4]. In order to demonstrate the clinical value of the proposed method, we compared the asymmetry index estimated using NABS and ADisc in normal aging and AD. Based on this index, we analyzed the population detection capability for both methods.

For this purpose, we used T1-weighted MR images of 40 right-handed subjects from the OASIS dataset [37] between 60 and 90 years old (a control group of 10 males and 10 females considered non-demented (Clinical dementia rating (CDR) =0) and a group of 10 males and 10 females diagnosed to suffer from mild-dementia (CDR=1). We processed the images of the subjects with the two compared segmentation methods and calculated the asymmetry for cerebrum and cerebellum (eq. 10). Finally, we analyzed the differences between non-demented and mild-

demented groups. We studied the absolute values of the asymmetry index to better detect the asymmetry rather than the lateralization of brain atrophy [38]. One should note that we used the same template library of 30 MRIs of normal subjects for NABS segmentations as used in the previous experiments. In other words, constructing an AD specific training library was not required.

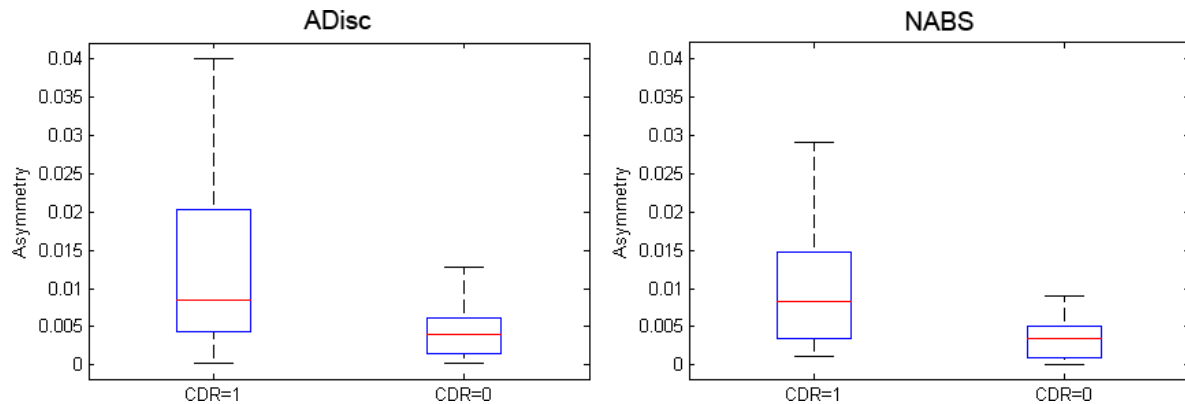


Figure 11. Box Whisker representation of the CH asymmetry coefficients of CDR=1 (mild dementia) and CDR=0 (cognitively normal) subjects for ADisc and NABS methods.

We evaluated the results by calculating a p-value using a t-test. No significant asymmetry differences were found between male and female subjects within each group. We found significant differences in cerebrum asymmetry between control and mild-demented groups that is in good agreement with previously published results [4] with both methods, p-values of 0.0012 (t-statistic 3.51, degrees of freedom 38) for ADisc and 0.0011 (t-statistic 3.51, degrees of freedom 38) for NABS, as shown in figure 11). We found significant differences for cerebellum asymmetry between mild-demented and the normal control populations with NABS segmentation method ($p = 0.028$, t-statistic = 2.28, degrees of freedom 38) but not with Adisc ($p = 0.5257$, $t = 0.604$, degrees of freedom 38). To further investigate the detected differences between groups, we computed the receiver operating characteristic (ROC) curves for estimated asymmetry as a predictor of mild-dementia (these are shown in Figure 12) and we estimated the areas under ROC curves (AUCs) as indexes summarizing the information contained in the ROC curves [39]. The AUC for NABS asymmetries were 0.80 for cerebrum and 0.70 for cerebellum and AUC for ADisc asymmetry was 0.77 for cerebrum and 0.53 for cerebellum. DeLong's test [40] did not indicate significant differences between AUCs provided by both segmentation methods. This may be because of the small sample size. However, the ROC curves in Figure 12 suggest that NABS

asymmetries were better predictors of AD.

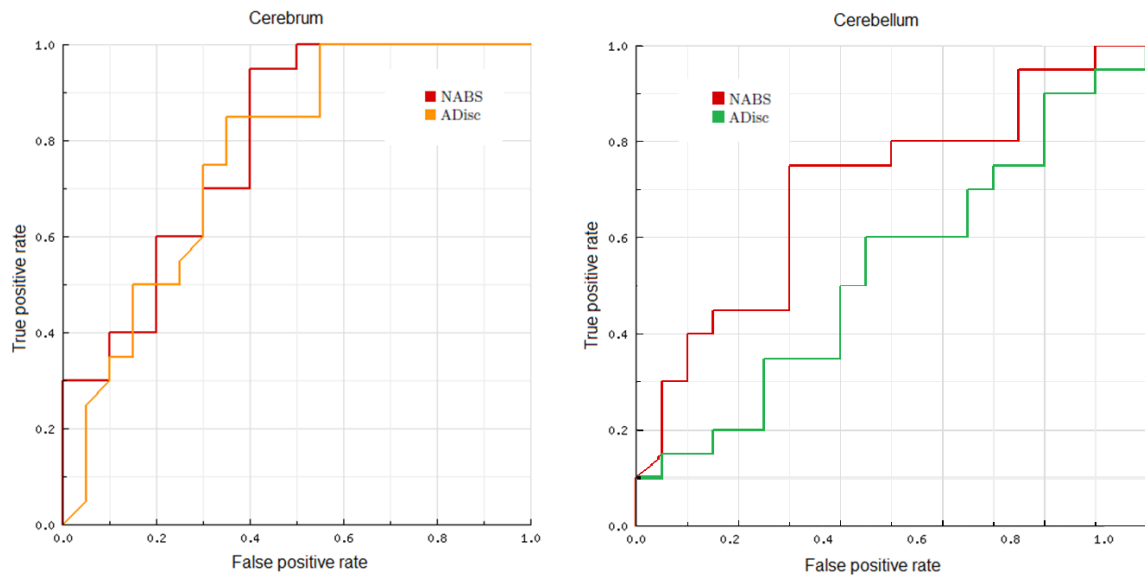


Figure 12. ROC curves comparison for NABS and ADisc results. Left graph stands for cerebrum and right graph for cerebellum.

To illustrate this, in figure 13, one can observe a similar hemisphere boundary alteration to the one shown in figure 10 for the ADisc method. This might be caused by a significant volume difference between both hemispheres. It is worth noting that the higher asymmetry difference, while not significant probably owing to the small sample size, found between MCI subjects and normal controls by NABS highlights the potential of this method for clinical studies. Its robustness and accuracy could enable the detection of finer anatomy alterations caused by pathologies.

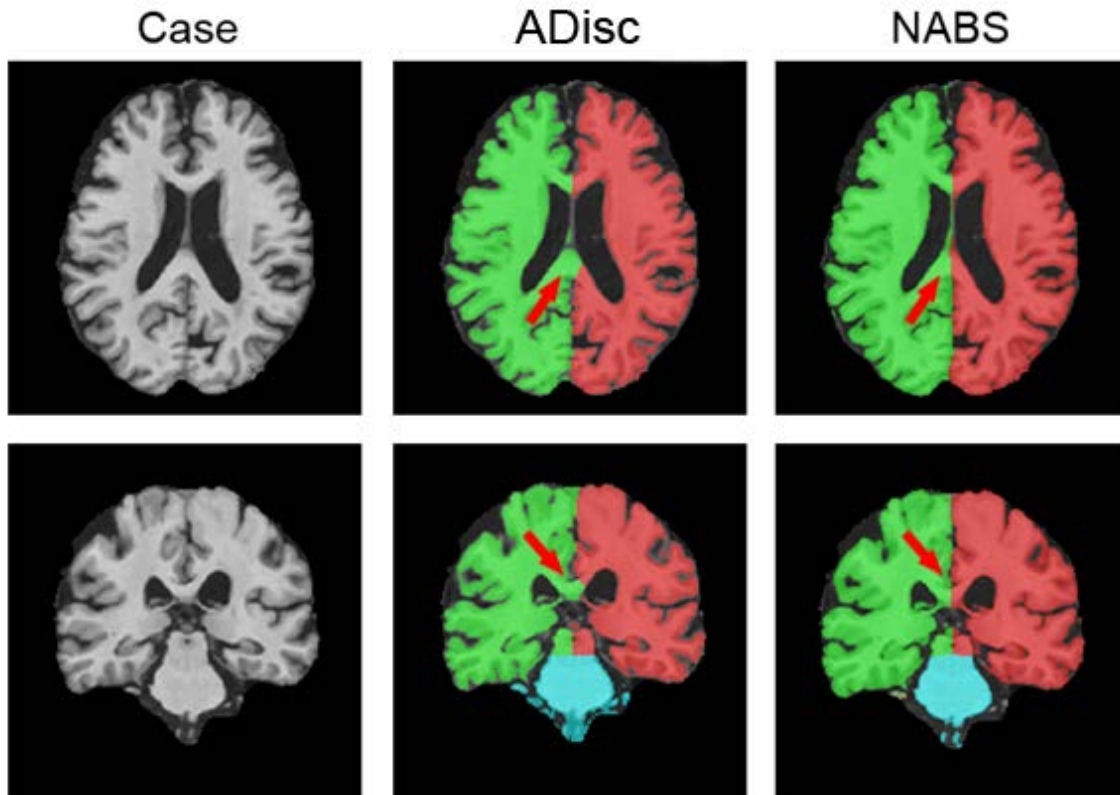


Figure 13. From left to right: Pre-processed image, ADisc and NABS segmentation results for a mild dementia affected case. The red arrows point to locations where ADisc method failed while NABS produced a correct classification.

We also analyzed the relations between the Mini Mental State Exam (MMSE) score and the degree of brain asymmetries. We found a significant correlation between MMSE and the degree of brain asymmetry in cerebrum with both segmentation methods. However, with NABS, the correlation was greater than with ADisc (ADisc correlation coefficient was -0.54 ($p = 0.0003$, 95% confidence interval from -0.73 to -0.27) while NABS correlation coefficient was -0.68 ($p = 0.00004$, 95% confidence interval from -0.83 to -0.46)).

3.3.5. Alzheimer's disease case

A final experiment was conducted to illustrate the robustness of NABS method to the presence of extreme brain abnormalities/atrophies. An OASIS database case with a large amount of brain atrophy was processed with the compared methods. The subject was an 86 years old male

affected by the Alzheimer's disease (CDR=2). As can be noticed in figure 14, NABS method provided acceptable results while ADisc method mislabels several regions (especially cerebellum).

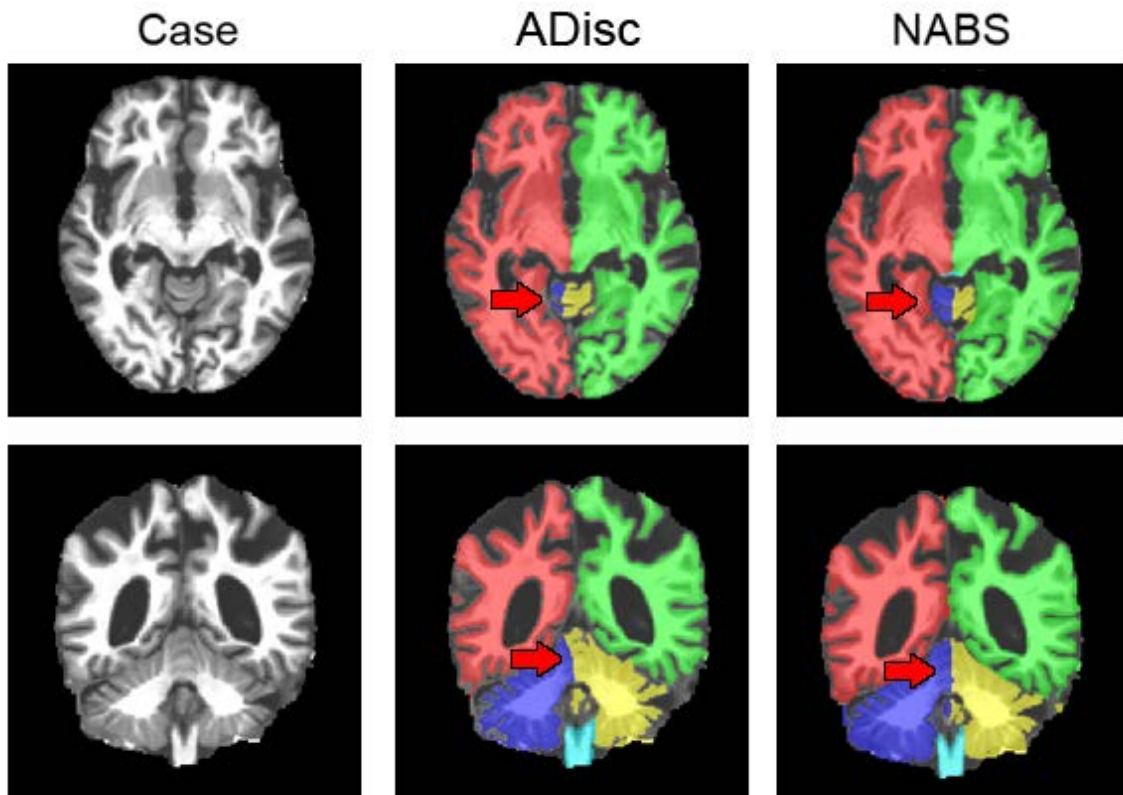


Figure 14. From left to right: Pre processed image, ADisc and NABS segmentations. The red arrows highlight the areas where obvious errors were found in ADisc result while NABS generated a correct labeling.

3.3.6. Computation time

ADisc segmentation required on average 5 minutes using the optimized C-based implementation reported [Zhao and Tohka, 41] which markedly improved the reported running time of approximately 1 hour of the original version). NABS segmentations required approximately 60 seconds on average. NABS was implemented as a multithreaded MEX C file. All timings were performed on Intel core i7 CPUs with 8 GB RAM running MATLAB 2009b 64 bits under Windows 7 Professional.

4. Discussion

Previous studies [23, 25] have shown that non-local patch-based label fusion approaches using a library of expert priors is a powerful approach for brain structure segmentation. In this paper, we have presented an accurate and fast patch-based multi-template brain segmentation method, termed NABS, for segmenting cerebral and cerebellar hemispheres and brainstem from T1-weighted MR brain images. The main methodological novelties of NABS over other similar methods are the use of a multi-label block-wise label fusion strategy specifically designed to deal with the classification of large brain structure and the use of an optimized mask for restricting the voting to only specific parts of the brain images. These optimizations drastically reduced the required computation time to approximately 1 minute per case.

We have shown that NABS method was able to accurately delineate brain structures in healthy subjects across a wide range of ages. Furthermore, we have presented evidence that suggests that NABS was still accurate even with highly atrophic and lesioned brains even when using a template library containing only images from healthy subjects from a different database. This is notable because it indicates that it is not necessary to construct disease specific templates to use NABS. Obviously, it is still not completely known how this new methodology will be affected when processing cases from different pathologies with significant anatomical alterations. However, although the method could fail in some non-typical cases, these problematic cases can be manually segmented and added to the library in order to improve the robustness of the system.

We have also provided quantitative comparisons against the ADisc method [11], which represents the state-of-art for brain hemisphere segmentation and has been validated against several hemisphere segmentation methods. These comparisons demonstrated that the NABS was consistently and significantly more accurate than ADisc. As a practical application, we studied the brain asymmetry indexes in MCI based on the segmentation provided by both NABS and ADisc. Both methods detected significant asymmetry increase in the MCI as compared to age and gender matched normal controls what is in accordance with previous literature [4]. The effect sizes were larger (and p-values were smaller) with NABS than with ADisc, which could be indicative of a more accurate segmentation by NABS. In addition, the MMSE test results were found to correlate between asymmetry index, again with a larger correlation coefficient based on the segmentations by NABS.

We attribute the success of the proposed method to the high degree of inter-subject pattern redundancy near the inter-hemispheric boundary. In other words, given an image to be segmented it is highly likely that there exists an image in the template library which has very similar inter-hemispheric anatomical patterns.

5. Acknowledgments

We want to thank the OASIS (P50 AG05681, P01 AG03991, R01 AG021910, P50 MH071616, U24 RR021382, R01 MH56584) and IXI - Information eXtraction from Images (EPSRC GR/S21533/02) datasets promoters for making available this valuable resource to the scientific community which surely will boost the research in brain imaging. This work has been supported by the Spanish grant TIN2011-26727 from Ministerio de Ciencia e Innovación. J. Tohka's work was supported by the Academy of Finland grant 130275.

References

- [1] Toga W and Thompson PM. Mapping brain asymmetry. *Nature Reviews Neuroscience* 2003;4:37-48.
- [2] Nielsen JA, Zielinski BA, Ferguson MA, Lainhart JE, Anderson JS. An Evaluation of the Left-Brain vs. Right-Brain Hypothesis with Resting State Functional Connectivity. *Magnetic Resonance Imaging. PLoS ONE* 2013;8(8): e71275. doi:10.1371/journal.pone.0071275.
- [3] Harrison PJ. The neuropathology of schizophrenia. A critical review of the data and their interpretation. *Brain* 1999;122(4):593-624.
- [4] Thompson PM, Moussai J, Zohoori S, Goldkorn A, Khan AA, Mega MS, et al. Cortical variability and asymmetry in normal aging and Alzheimer's disease. *Cerebral Cortex* 1998;8:492-509.
- [5] Herbert MR, Ziegler DA, Deutsch CK, O'Brien LM, Kennedy DN, Filipek PA, et al. Brain asymmetries in autism and developmental language disorder: a nested whole-brain analysis. *Brain* 2005;128:213-26.
- [6] Bilder RM, Wu H, Bogerts B, Degreef G, Ashtari M, Alvir JM, et al. Absence of regional hemispheric volume asymmetries in first-episode schizophrenia. *Am J Psychiatry* 1994;151:1437-47.
- [7] Pepe A, Zhao L, Koikkalainen J, Hietala J, Ruotsalainen U and Tohka J. Automatic statistical shape analysis of cerebral asymmetry in 3D T1-weighted magnetic resonance images at vertex-level: Application to neuroleptic-naïve schizophrenia. *Magnetic Resonance Imaging* 2013;doi: 10.1016/j.mri.2012.10.021.
- [8] Liang L, Rely K, Woods R and Rottenberg D. Automatic segmentation of left and right cerebral hemispheres from MRI brain volumes using the graph cuts algorithm. *NeuroImage* 2007;34:1160-1170.
- [9] Zilles K, Kawashima R, Dabringhaus A, Fukuda H, Schormann T. Hemispheric shape of European and Japanese brains: 3-D MRI analysis of intersubject variability, ethnical, and gender differences. *Neuroimage* 2001;13(2):262-71.

- [10] Zhao L, Hietala J and Tohka J. Shape analysis of human brain interhemispheric fissure bending in MRI. *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2009, Lecture Notes in Computer Science* 2009;5762:216 - 223.
- [11] Zhao L, Ruotsalainen U, Hirvonen J, Hietala J and Tohka J. Automatic cerebral and cerebellar hemisphere segmentation in 3D MRI: adaptive disconnection algorithm. *Medical Image Analysis* 2010;14: 360-372.
- [12] Brummer M. Hough transform detection of the longitudinal fissure in tomographic head images. *IEEE Transactions on Medical Imaging* 1991;10: 74–81.
- [13] Sun C and Sherrah J. 3D symmetry detection using the extended Gaussian image. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1997;19: 164-168.
- [14] Dale A, Fischl B and Sereno M. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage* 1999;9: 179–194.
- [15] Krigeskorte N and Goebel R. An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical MR volumes. *NeuroImage* 2001;14: 329-346.
- [16] Marais P and Brady J. Detecting the brain surface in sparse MRI using boundary models. *Medical Image Analysis* 2000;4:283-302.
- [17] Maes F, Van Leemput K, Delisi L, Vandermeulen D and Suetens P. Quantification of cerebral grey and white matter asymmetry from MRI. *Lecture Notes in Computer Science* 1999;1679:348–357.
- [18] Larsson J. *Imaging vision: functional mapping of intermediate visual processes in man*. Ph.D. Thesis, Karolinska Institutet, Stockholm, Sweden. 2001;ISBN 91-7349-090-3.
- [19] Mangin JF., Rivière D, Cachia A, Duchesnay E, Cointepas Y, Papadopoulos-Orfanos D, et al. A framework to study the cortical folding patterns. *NeuroImage* 2004;23:129–138.
- [20] Hata Y, Kobashi S, Hirano S, Kitagaki H and Mori E Automated segmentation of human brain MR images aided by fuzzy information granulation and fuzzy inference. *IEEE Transactions on Systems, Man and Cybernetics Part C – Applications and Reviews* 2000;30:381-395.

- [21] Collins DL and Pruessner JC. Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion. *NeuroImage* 2010;52(4):1355-66.
- [22] Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage* 2009;46:726-738.
- [23] Coupé P, Manjón JV, Fonov V, Pruessner J, Robles M and Collins DL. Patch-based Segmentation using Expert Priors: Application to Hippocampus and Ventricle Segmentation. *NeuroImage* 2011;54:940-954.
- [24] Wang H, Suh JW, Das S, Pluta J, Craige C and Yushkevich P, Multi-Atlas Segmentation with Joint Label Fusion. *PAMI* 2013;35(3):611-623.
- [25] Eskildsen SF, Coupé P, Fonov V, Manjón JV, Leung KK, Guizard N, et al. BEaST: Brain extraction based on nonlocal segmentation technique. *Neuroimage* 2012;59:2362-2373.
- [26] Rousseau F, Habas P, Studholme C. A supervised patch-based approach for human brain labeling. *IEEE Transactions on Medical Imaging* 2011;30:1852-1862.
- [27] Manjón JV, Coupé P, Martí-Bonmatí L, Robles M and Collins L. Adaptive Non-Local Means Denoising of MR Images with Spatially Varying Noise Levels. *Journal of Magnetic Resonance Imaging* 2010;31:192-203.
- [28] Penny WD, Friston KJ, Ashburner JT, Kiebel SJ and Nichols TE. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press 2007;SBN: 978-0-12-372560-8.
- [9] Manjón JV, Tohka J and Robles M. Improved Estimates of Partial Volume Coefficients from Noisy Brain MRI Using Spatial Context. *Neuroimage* 2010;53:480-490.
- [30] Lotjonen J, Wolz R, Koikkalainen J, Thurfjell L, Waldemar G, Soininen H, et al. Fast and robust multi-atlas segmentation of brain magnetic resonance images. *Neuroimage* 2010;49(10):2352-2365.
- [31] Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* 2006;31:1116-28.

- [32] Wang Z, Bovik AC, Sheikh HR and Simoncelli EP. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing* 2004;13:600-612.
- [33] Coupé P, Yger P, Prima S, Hellier P, Kervrann C and Barillot C An Optimized Blockwise NonLocal Means Denoising Filter for 3-D Magnetic Resonance Images. *IEEE Transactions on Medical Imaging* 2008;27:425-441.
- [34] Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurements* 1960;20:37-46.
- [35] Mangin, JF, Régis J, and Frouin V. Shape bottlenecks and conservative flow systems. *MMBIA Proceedings* 1996;319-328.
- [36] Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE Trans. Med. Imaging* 1994;13:716-724.
- [37] Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL. Open Access Series of Imaging Studies (OASIS): Cross-Sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *Journal of Cognitive Neuroscience* 2007;19:1498-1507.
- [38] Derflinger S, Sorg C, Gaser C, Myers N, Arsic M, Kurz A, et al. Grey-matter atrophy in Alzheimer's disease is asymmetric but not lateralized. *J Alzheimers Dis* 2011;25(2):347-57.
- [39] Hanley JA, McNeil BJ. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology* 1982;143(1):29-36.
- [40] Vergara IA, Norambuena T, Ferrada E, Slater AW and Melo F. StAR: a simple tool for the statistical comparison of ROC curves. *BMC Bioinformatics* 2008;9:265.
- [41] Zhao L. and Tohka J. ADisc - A pipeline for Adaptive Disconnection based brain hemisphere segmentation in 3D MRI. 17th Annual Meeting of the Organisation for Human Brain Mapping Quebec City, Canada. 2011.