Applied Medical Informatics
Original Research Vol. 37, No. 4 /2015, pp: 1-10

A Multiatlas Approach for Segmenting Subcortical Brain Structures using Local Patch Distance

Neela RAMAMOORTHI[1,*] and Kalaimagal RAMANUJAM[2]

[1]Research Scholar, Manonmaniam Sundaranar University, Tirunelveli-627 012, Tamil Nadu, India
[2]Department of Computer Science, Govt. Arts College for Men (Autonomous), Nandhanam, Chennai - 600 035. Tamilnadu, India
E-mails: avcneela@gmail.com, kalai_magal@rediffmail.com

* Author to whom correspondence should be addressed; Tel.: +919788596185

Received: June 14, 2015 /Accepted: November 28, 2015/ Published online: December 15, 2015

Abstract
In the diagnosis and treatment of various diseases, often segmenting the brain structures from MRI data is the key step. Since there are larger variations in the anatomical structures of the brain, segmentation becomes a crucial process. Using only the intensity information is not enough to segment structures since two or more structures may share the same tissues. Recently, the use of multiple pre-labeled images called atlases or templates are used in the process of segmentation of image data. Both single atlas and multiple atlases can be used. However, using multiple atlases in the segmentation process proves a dominant method in segmenting brain structures with challenging and overlapping structures. In this paper, we propose two multi atlas segmentation methods: Local Patch Distance Segmentation (LPDS) and Weighted Local Patch Distance Segmentation (WLPDS). These methods use local patch distance in the label fusion step. LPDS uses local patch distance to find the best patch match for label propagation. WLPDS uses local patch distance to calculate local weights. The brain MRI images from the MICCAI 2012 segmentation challenge are chosen for experimental purposes. These datasets are publicly available and can be downloaded from MIDAS. The proposed techniques are compared with existing fusion methods such as majority voting and weighted majority voting using the similarity measures such as Dice overlap (DC), Jaccard coefficient (JC) and Kappa statistics. For 20 test data sets, LPDS gives DICE=0.95±0.05, JACCARD=0.91±0.04 and KAPPA=0.94±0.07. WLPDS gives DICE=0.98±0.02, JACCARD=0.92±0.03 and KAPPA=0.95±0.04.

Keywords: Brain MRI images; Subcortical Structures; Local Patch Distance Segmentation (LPDS); Weighted Local Patch Distance Segmentation (WLPDS)

Introduction

Segmenting brain tissues/ structures is the key step in the diagnosis and treatment of various brain related disorders. The MRI scan gives clear and exhaustive images of soft tissue. Bones can't be visualized using MRI because bone tissue contains small amounts of water. Due to this, MRI and fMRI are widely used in tissue and structure segmentation. Tissue segmentation can be done using intensity threshold. But segmenting brain structures is complicated as a structure may be comprised of more than one tissue type. Because of this, segmenting brain structures represents a very crucial step. Though a lot of research has been done in this area, still it remains a challenging field. Using prior knowledge about the spatial relationships among structures, called atlases, we can

segment the structures with larger dissimilarities. Among other proposed methods, atlas based methods give best result [1]. Multiple atlases [2,3] prove to be the best method when compared to single atlas, especially when there are larger dissimilarities in the structures. If a single atlas is used during segmentation, then the result will be biased. To remove such biases and also segment structures with larger dissimilarities, multiple atlases are used. When multiple atlases are used the accuracy of the segmentation is not that much affected by the errors produced during registration. This method has two steps, registration and fusion. The registration is done for each atlas image with the test image to align both of them into a common space and the labels from the atlas image are propagated to the test image. This step gives a set of segmentations, one for each atlas image. Then the fusion of labels from all segmentations is done to obtain the final true segmentation.

An atlas is a set of intensities and their segmented label image. The accuracy of the segmentation result largely depends on the selection of atlases [4]. Various techniques can be applied in the label fusion step. Majority voting [4-6] is considered as the simplest label fusion strategy and is widely used one. Weighted voting [7] is a technique which assigns a positive weight, calculated using the atlas target similarities (similarities can be globally [2,6], locally [7,8] or combined [9]). The patch based method proposed by Coupe et al. [10]. Kittler et al. [11] gives an overview of various fusion rules. An expectation maximization method is used in the STAPLE algorithm to compute weight [12]. The segmentation pipeline using multiple atlases is shown in Figure 1. A new method for selecting a best matching patch in the label fusion step, called Local Patch Distance Segmentation, which computes a local patch distance using neighborhood intensity similarities is proposed. Also a variation in LPDS called WLPDS is developed for calculating local weights.

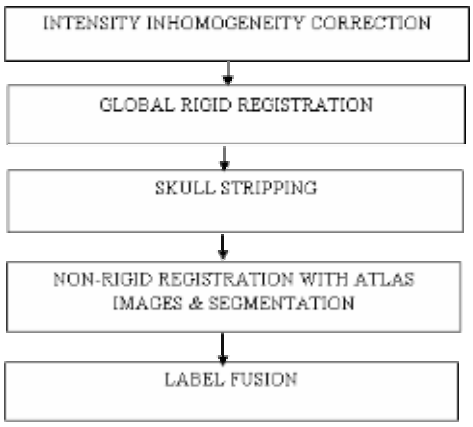

Figure 1. Segmentation pipeline

Materials and Methods

Let us denote the image to be segmented by I_T and atlas image set by A={A1,A2,…AN} where there are 'N' atlases. An atlas, Ai, is an intensity and label pair denoted by AIi and ALi respectively. The intensity heterogeneity is caused mainly due to the RF coil imperfections [13,14]. This artifact results in the variation of intensities in the same tissues. These variations affect the segmentation results. The first intensity heterogeneity correction should be done on all images. Then both test and atlas images are aligned rigidly to a common template for further processing. The brain tissues should be extracted from the non-brain area via a process called skull stripping for good results and to increase the speed of the algorithm. Then the test image I_T should be non-rigidly aligned with each atlases Ai which yields a transformation field for each registration. Using this transformation field the labels from corresponding atlas label image Ali are propagated to produce an equivalent intermediate segmentation Si. This registration process yields 'N' intermediate segmentations, Si, where i=1,2,…N, one for each atlas. Our objective is to combine all these intermediate segmentations and produce the final true segmentation, T_{Final}, using some

combination rule. The segmentation accuracy is computed by comparing T_{Final} with the golden standard (provided in the training set) using any similarity measures. Our problem is to formulate a fusion approach that produces T_{Final} with greater accuracy. Our algorithm should segment multiple structures even though the intensity of the atlas and target images varies highly. For each voxel, we consider both the performance of the atlas as well as the spatial information in addition to the intensity during the labeling process. During label fusion process, the STAPLE [12] method considers the performance of the atlas whereas the simple voting method does not. Both the STAPLE [12] and simple voting methods do not consider the intensity information in the labeling process.

Data

ADNI, LONI, LPBA 40, SATA and IXI are some of the publicly available data sets that can be used for the evaluation of segmentation algorithms. SATA(MICCAI 2012 Challenge on Segmentation: Algorithms, Theory and Applications) data set was used for evaluation of the proposed algorithm. There are 15 labeled datasets and 20 unlabeled data sets, which are bias corrected. These datasets are downloaded from MIDAS digital archive system [15]. We segmented the brain structures for all the images by selecting one image as target image and remaining as atlases from the data set using leave-one-out strategy.

Local Patch Distance Segmentation (LPDS)

In this paper, we aimed to design a system that automatically segments T1-weighted brain MRI volumes into anatomical sub-cortical and cortical structures. We approached the 3D brain image segmentation problem as a minimization problem. We also proposed a variation of our method, using local patch distance to find local weights and then using majority voting rule to select the target labels. Patch based method operates on the local voxel's intensity values of an image. Figure 2 shows the framework for LPDS. From the atlas set, a set of atlases are selected based on the following criteria:

- Find the normalized mutual information between each image from the template and the test image
- The closer the value the more the images will be similar.
- Select the top 10 most used images as atlases [4].

Usually most of the label fusion step catches the similarities among the neighborhood voxels. A center voxel and its neighborhood voxels in certain shape form a patch. Usually the selected shape will be cubical. For each voxel in an image, a patch, a vector of size $m \times 1$ is considered. The test image I_T is divided into n patches (PTi, i=1…n) and the N atlases are divided into A_k patches (k=number of patches n * number of atlases N). Let PT ={PT_i} and PA ={A_k} be two sets, test and atlases patches respectively. Our objective is to find a best match for every patch PT in PA and assign the label of PA (center voxel) to the center voxel in PT. Usually the patch based methods does not require registration. Here, in our approach, for greater accuracy, we non-linearly registered each atlas and test image separately for matching purposes. Our method uses local patch distance to measure similarity for selecting the best matching patch. It is calculated for each patch from atlas images and from target test image. For each patch from the target image, the patch with the smallest local patch distance gives a best matching patch from the atlas image. The label from that best match patch (label of center voxel) is transferred to the center voxel in the target patch. These steps are repeated for all the patches in the target image. The algorithm for calculating the local patch distance is given below:

- Divide the test image (I_T) and all atlases (A_i i=1, ..N) into equal sized patches
- For each patch in the test image, PTj , and for every atlas, PA_i repeat the following steps:
  - Start search for the similar matching patch in an atlas image from the same location.
  - If a match is found then set local patch distance as 0 (means this the best match)

- If not, then search in the neighborhood patches of that atlas, calculate the offset between these two patches and sum up the offset. Repeat this step until a matching patch is found or until the maximum neighborhood distance (parameter) is reached.

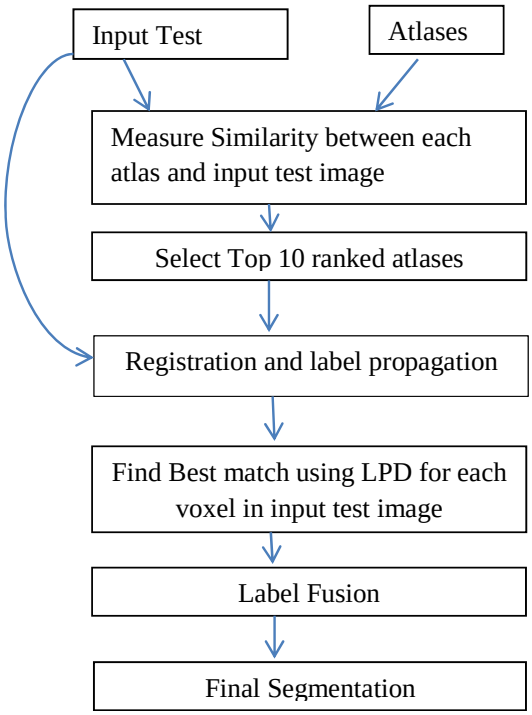

Figure 2. Local Patch Distance Segmentation Framework

Weighted Local Patch Distance Segmentation (WLPDS)

In WLPDS, after calculating the local patch distance for the target image, weights will be calculated using the weighted scheme. More weights (using 1) will be assigned to the patch that has the smallest local patch distance.

$$w(PTi,PAi)=\frac{e^{-\frac{d(PTi,PAi)}{z^2}}}{\sum_{x\in X} e^{-\frac{d(PTi,x)}{z^2}}} \quad (1)$$

where d(PTi, PAi) are the local patch distance between patches PTi, PAi, z>0 smoothing parameter.

Then after assigning weights, the labels are fused using weighted majority voting rule to form the final segmentation, T_{Final}.

Dice coefficient, Jaccard Coefficient, False positive rate, False negative rate, and Kappa statistics as metrics were used for evaluation of the proposed method. The segmentation results were compared with the pre-labeled image from the training set provided by MIDAS. The dice overlap is calculated by using

$$D(A,B)=[2|A\cap B|]/[|A|+|B|] \quad (2)$$

Jaccard coefficient is calculated by using

$$J(A,B)=(A\cap B)/(A\cup B) \quad (3)$$

where A and B refers to final segmentation and ground truth image respectively; the value 0 gives dissimilarity and 1 gives similarity.

Segmentation errors are calculated as false positive (FP) and false negative (FN) using

$$FP=|A|B|/B \quad (4)$$

$$FN=|B|A|/B \quad (5)$$

where A and B refers to the final segmentation and ground truth image respectively

Since all the images provided in MIDAS are already homogeneity corrected, we omit this preprocessing step. However for bias correction we use ANTs N4 bias correction algorithm described by Tustison et al. [16]. After bias correction, all images (both test and atlases) are rigidly aligned with MNI template space using 3D Slicer's BrainsFit [17] for further processing. For removing non-brain tissues, the ROBEX [18] skull stripping algorithm proposed by Iglesias et. al. [19] was used. The top 10 atlases [4] are selected from the atlas set by calculating similarity, using mutual information between the test image and each atlas. Then the test image is non-rigidly registered with each atlas image and the labels from the atlas image are propagated to the aligned test image using ANTs tool [20]. This gives 10 different segmentations, one for each atlas that are fused using our proposed methods. Figure 3 (a) to (e) shows the result of segmentation of a sample test image for one atlas.

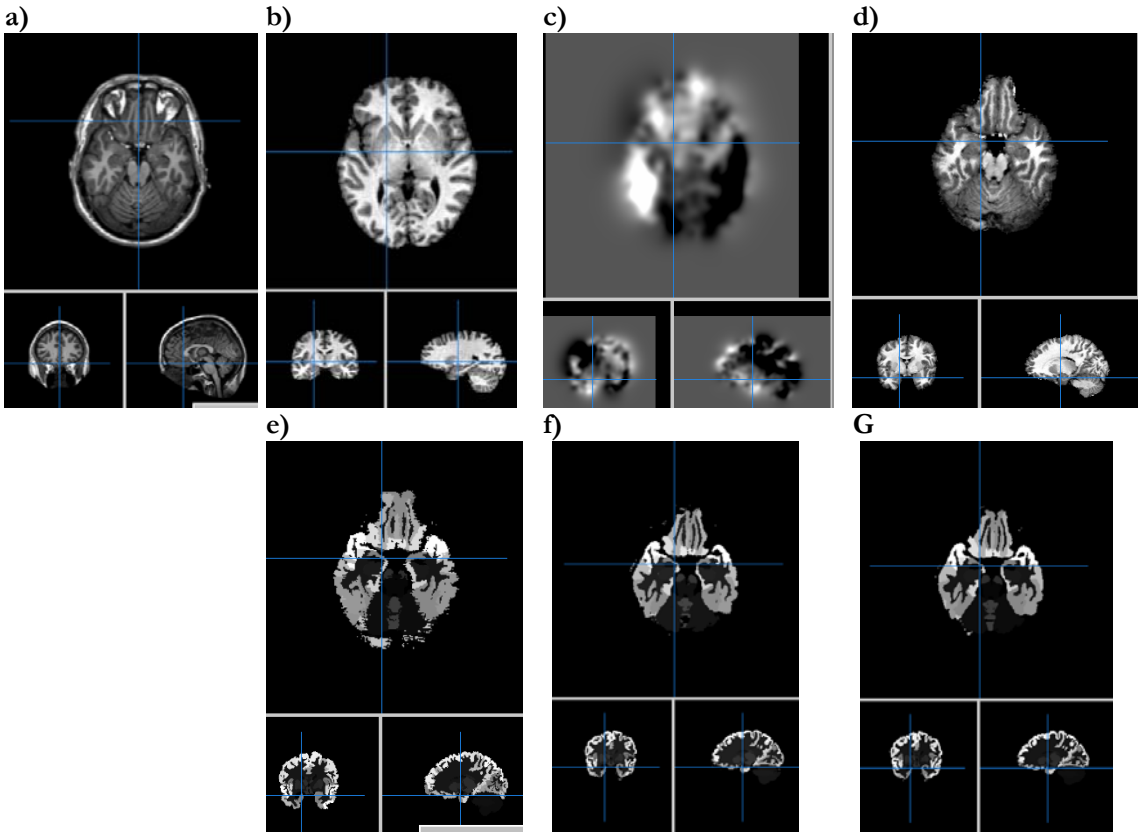

Figure 3. Segmentation result for one test image using WLPDS – (a) Test Image; (b) Test image after skull stripping using ROBEX; (c) Test image after non-rigid registration using ANTs; (d) Segmentation of test image; (e) Propagation of Label from atlas to test image; (f) T_{Final} after of (e) for all remaining atlases using LPDS; (g) T_{Final} after of (e) for all remaining atlases using WLPDS

The Final segmentation after fusion of all intermediate segmentations obtained by registration of each atlas with the test image is shown in Gigure 3(g) and 3(h). Majority voting and weighted majority voting was in comparison of the proposed approach.

Our experiment was conducted for all the images in MICCAI 2012 challenge data set. Our algorithm was tested on an Intel® Core ™ i3-3271U CPU @ 1.80 GZ using ITK [21]. Among the operations, only non-rigid registration and fusion steps consume more time. For doing each set of

registration, 30 to 40 minutes are needed. To do the final fusion, the proposed algorithm takes 1 ½ hours. The time taken to compare the result with the golden standard was 49332 milliseconds.

Results

Table 1 and Figure 4 show the performance comparison of our methods with simple majority voting and weighted majority voting in terms of Dice, Jaccard, and Kappa statistics.

Table 1. Average segmentation performance comparison of different label fusion strategies - LPDS, WLPDS, MV and WMV for 20 training data sets in terms of Dice, Jaccard and Kappa statistics(±standard deviation).

Methods / Measures	LPDS	WLPDS	MV	WMV
DICE	0.9547±0.05	0.9854±0.02	0.8245±0.95	0.9245±0.39
JACCARD	0.9133±0.04	0.9231±0.03	0.9021±2.01	0.9453±1.10
KAPPA	0.9453±0.07	0.9544±0.04	0.9321±2.98	0.8934±1.03

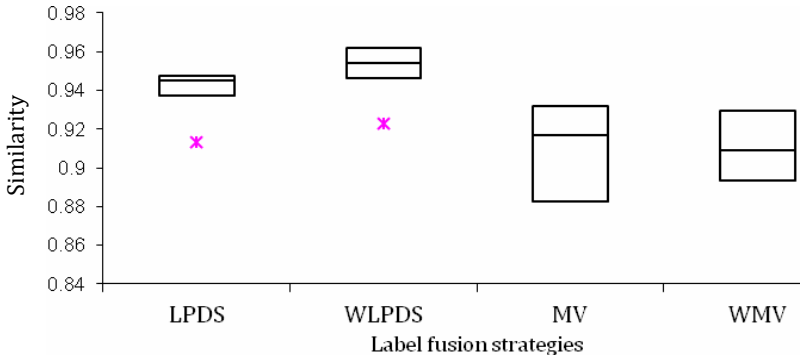

Figure 4. Comparison of segmentation performance of LPDS, WLPDS, MV & WMV in terms of Dice Similarity, Jaccard coefficient and Kappa Statistics

Because the overall performance of WLPDS is greater when compare to LPDS, the performance of segmentation of various structures for WLPDS were compared and the results are given in Table 2 & Figure 5.

Table 2. The average dice overlap, Jaccard, Target, False Negative and False Positive for segmenting various structures for 20 data sets using WLPDS

Name	Target	Jaccard	Dice	FN	FP
CSF	0.9697	0.8397	0.9129	0.0304	0.1376
Left Amygdala	0.9648	0.8852	0.9391	0.0352	0.0852
Left Hippocampus	0.9553	0.9088	0.9522	0.0447	0.0508
Left Caudate	0.9543	0.9251	0.9611	0.0457	0.0319
Left Pallidum	0.9529	0.9196	0.9581	0.0471	0.0366
Left Putamen	0.9774	0.9433	0.9708	0.0226	0.0357
Right Amygdala	0.9235	0.8729	0.9321	0.0766	0.0590
Right Hippocampus	0.9402	0.8960	0.9452	0.0598	0.0498
Right Caudate	0.9647	0.9303	0.9639	0.0353	0.0369
Right Pallidum	0.9340	0.9008	0.9478	0.0661	0.0379
Right Putamen	0.9802	0.9594	0.9793	0.0198	0.0217

FN = false negative; FP = false positive

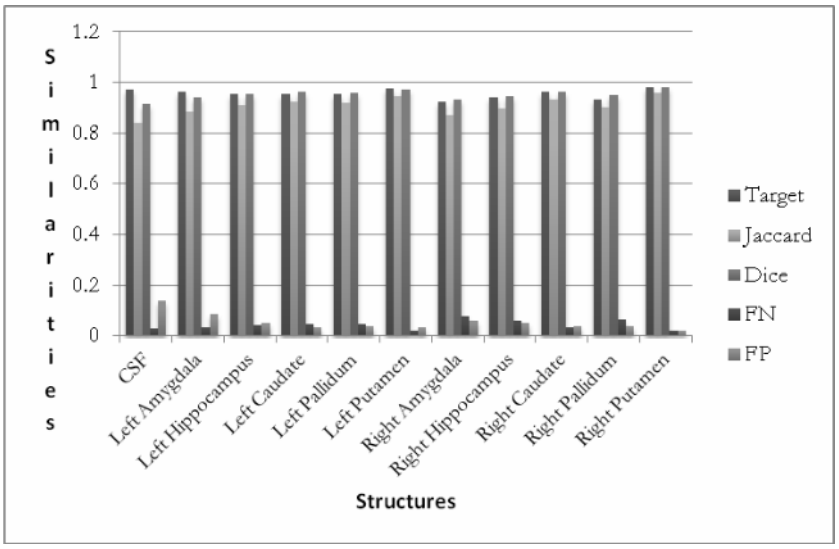

Figure 5. Evaluation of WLPDS for the test image in Figure 3a

Table 3 and Figure 6 and show the segmentation performance of WLPDS for various numbers of atlases.

Table 3. Segmentation performance comparison of WLPDS for the test image (figure 3a) for a different number of atlases

Atlases	Total	JC	DC	Volume sim.	False negative	False positive
10	0.8957	0.8225	0.9026	-0.0155	0.1043	0.0904
7	0.8801	0.8020	0.8902	-0.0228	0.1199	0.0996
5	0.8536	0.7726	0.8717	-0.0425	0.1464	0.1094
4	0.8171	0.7296	0.8436	-0.0651	0.1829	0.1280

JC = Jaccard; DC = Dice

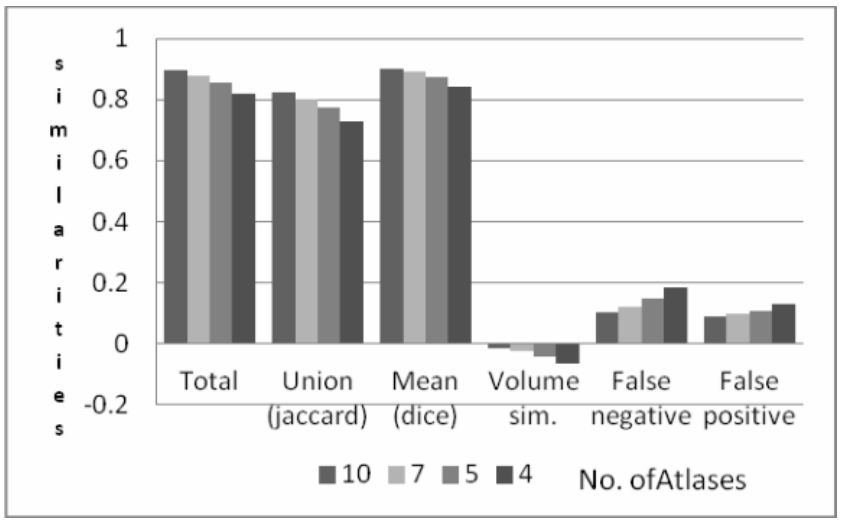

Figure 6. Comparison of WLPDS for different number of atlases

For evaluation of our methods, the average overlap and the standard deviation between the

segmented image and the golden standard image were compared. For the test image shown in Figure 3a, LPDS gives DICE=0.9547±0.05, JACCARD=0.9133±0.04 and KAPPA=0.9453±0.07. WLPDS gives DICE=0.9854±0.02, JACCARD=0.9231±0.03 and KAPPA=0.9543±0.04. The same test image gives DICE=0.8245±0.95, JACCARD=0.9021±2.01 and KAPPA=0.9321±2.98 for simple majority voting method and gives DICE=0.9245±0.39, JACCARD=0.9453±1.10 and KAPPA=0.8934±1.03 for weighted majority voting.

Discussion

Two variations in the label fusion step in the process of segmentation of brain image using multiple atlases were proposed in this manuscript. To label a voxel, the intensity and spatial information of neighboring voxels in addition to that particular voxel were used. The overall mean dice similarity for majority voting is 0.726±0.138, joint label fusion is 0.766±0.013 and joint label fusion with corrective learning [22], which won first place in the 2012 MICCAI Multi-Atlas Labeling Challenge is 0.782 ± 0.010. The overall mean dice similarity for LPDS is 0.732±0.136 and WLPDS is 0.739 ± 0.179, which are slightly greater than majority voting but lesser than joint label fusion and joint label fusion with corrective learning. The majority voting method assigns equal weights to every atlas. This doesn't consider the similarities between atlas and test image. Also the intensity information of the test image is also ignored [2]. Joint Label Fusion method considers the dependency among the test image and atlases in calculation of weights. The errors produced by registration are eliminated in this method. But errors can be produced by the spatial bias. To eliminate such bias, corrective learning technique is applied. An AdaBoost classifier was trained to correct this bias error. Our method takes both the spatial correspondence between the test image and atlas image in assigning the weights. The limitations of our methods are that the method is are time consuming and depend on the structure to be segmented. In future, the true patch based segmentation using local patch distance will be implemented which will ignore the registration between atlases and target image, the time consuming step. Hence the total time taken to segment the brain can be reduced and the speed of the segmentation process will be increased.

The key idea of our method is using the intensity information of surrounding voxels in addition to the current voxel to find the best match which is ignored by most of the current label fusion techniques. For weight calculation both intensity information as well as the atlas target matching information were used. This is the main advantage of our method. In our segmentation pipeline, the top 10 atlases were first selected using mutual information which also reduced the total number of registrations and hence reduces the total time taken for the entire segmentation process. Our further research is towards eliminating registration step and hence reduces the time taken for segmentation.

List of abbreviations

ADNI – Alzheimer's Disease Neuroimaging Initiative
ANTs – Advanced Normalization Tools
DC – Dice Overlap
fMRI – Functional MRI
FP – False Positive
FN – False Negative
IXI - Information eXtraction from Images
ITK –Insight Segmentation and Registration Toolkit
JC – Jaccard's Coefficient
KAPPA – Kappa Statistics
LONI – Laboratory of Neuro Imaging
LPBA 40 - LONI Probabilistic Brain Atlas
MRI – Magnetic Resonance Imaging

SATA - Segmentation: Algorithms, Theory and Applications

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Babalola K, Patenaude B, Aljabar P, Schnabel J, Kennedy D, Crum W, et al. An evaluation of four automatic methods of segmenting the subcortical structures in the brain. Neuroimage. 2009;47(4):1435-1447.
2. Heckemann R, Hajnal J, Aljabar P, Rueckert D, Hammers A. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. Neuroimage. 2006;33(1):115-126.
3. Artaechevarria X, Munoz-Barrutia A, Ortiz-de-Solorzano C. Combination Strategies in Multi-Atlas Image Segmentation: Application to Brain MR Data. IEEE Trans Med Imaging. 2009;28(8):1266-1277.
4. Aljabar P, Heckemann R, Hammers A, Hajnal J, Rueckert D. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. Neuroimage. 2009;46(3):726-738.
5. Rohlfing T, Brandt R, Menzel R, Maurer C. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. Neuroimage. 2004;21(4):1428-1442.
6. Sdika M. Combining atlas based segmentation and intensity classification with nearest neighbor transform and accuracy weighted vote. Medical Image Analysis. 2010;14(2):219-226.
7. Isgum I, Staring M, Rutten A, Prokop M, Viergever MA, Van Ginneken B. Multi-atlas-based segmentation with local decision fusion--application to cardiac and aortic segmentation in CT scans. IEEE Transactions on Medical Imaging 2009;28(7):1000-1010.
8. Wang H, Suh J, Das S, Pluta J, Craige C, Yushkevich P. Multi-Atlas Segmentation with Joint Label Fusion. IEEE Trans Pattern Anal Mach Intell. 2013;35(3):611-623.
9. Sabuncu M, Yeo B, Van Leemput K, Fischl B, Golland P. A Generative Model for Image Segmentation Based on Label Fusion. IEEE Trans Med Imaging. 2010;29(10):1714-1729.
10. Coupé P, Manjón J, Fonov V, Pruessner J, Robles M, Collins D. Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. Neuroimage. 2011;54(2):940-954.
11. Kittler J, Hatef M, Duin R, Matas J. On combining classifiers. IEEE Trans Pattern Anal Mach Intell. 1998;20(3):226-239.
12. Warfield S, Zou K, Wells W. Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation. IEEE Trans Med Imaging. 2004;23(7):903-921.
13. Hou Z, Huang S, Hu Q, Nowinski W. A fast and automatic method to correct intensity inhomogeneity in MR brain images. Med Image Comput Comput Assist Interv. 2006;9(2):324-331.
14. Vovk U, Pernus F, Likar B. A Review of Methods for Correction of Intensity Inhomogeneity in MRI. IEEE Transactions on Medical Imaging. 2007;26(3):405-421.
15. Placid.nlm.nih.gov. MIDAS - Digital Archiving System [Internet]. 2015 [cited 10June 2015]. Available from: http://placid.nlm.nih.gov/user/48
16. Tustison N, Avants B, Cook P, Yuanjie Zheng, Egan A, Yushkevich P et al. N4ITK: Improved N3 Bias Correction. IEEE Transactions on Medical Imaging. 2010;29(6):1310-1320.
17. Slicer.org. 3D Slicer [Internet]. 2015 [cited 10 June 2015]. Available from: http://www.slicer.org
18. Nitrc.org. NITRC: Robust Brain Extraction (ROBEX): Tool/Resource Info [Internet]. 2015 [cited 10 June 2015]. Available from: http://www.nitrc.org/projects/robex/

19. Iglesias J, Cheng-Yi Liu, Thompson P, Zhuowen Tu. Robust Brain Extraction Across Datasets and Comparison With Publicly Available Methods. IEEE Trans Med Imaging. 2011;30(9):1617-1634.
20. Stnava.github.io. ANTs by stnava [Internet]. 2015 [cited 10 June 2015]. Available from: http://stnava.github.io/ANTs/
21. ITK.ORG. ITK - Segmentation & Registration Toolkit [Internet]. 2015 [cited 10 June 2015]. Available from: http://ITK.org
22. Wang H, Yushkevich PA. Multi-Atlas Segmentation with Joint Label Fusion and Corrective Learning-an Open Source Implementation. Front Neuroinform. 2013;7:27.doi: 10.3389/fninf.2013.00027