Hippocampal segmentation with MAGeT

Jon Pipitone

December 13, 2012

Abstract

Neuroimaging research often relies on automated anatomical segmentations of MR images of the brain. Current multi-atlas based approaches provide accurate segmentations of brain images by propagating manually derived segmentations of specific neuroanatomical structures to unlabelled data. These approaches often rely on a large number of such manually segmented atlases that take significant time and expertise to produce. We present an algorithm for the automatic segmentation of the hippocampus that minimizes the number of atlases needed while still achieving similar accuracy to multi-atlas approaches.

Finish this...

1 Introduction

The hippocampus is of particular interest to many researchers because it is implicated in forms of brain dysfunction such as Alzheimer's disease and schizophrenia, and has functional significance in cognitive processes such as learning and memory. For many research questions involving magnetic resonance imaging (MRI) data accurate identification of the hippocampus and its subregions is a necessary first step to better understand the individual neuroanatomy of subjects.

Currently, the gold standard for neuroanatomical segmentation is manual delineation by an expert human rater. This is problematic for segmentation of the hippocampus for several reasons. First, manual segmentation takes a significant investment of time and expertise [7] which may not be readily available to researchers or clinicians. Second, the amount of data produced in neuroimaging experiments increasingly exceeds the capacity for identification of specific neuroanatomical structures by an expert manual rater. Third, the true definition of hippocampal anatomy in MR images is disputed [6], as evidenced by efforts to create an unified segmentation protocol [10].

Compounding each of these problems is the significant neuroanatomical variability in the hippocampus throughout the course of aging, development, and neuropsychiatric disorders [14]. Additionally, it may be necessary to use several different hippocampal definitions or, in fact, make specific modifications in the course of research. For example, Poppenk et al. [15] found that subdividing the hippocampus into anterior and posterior regions resulted in a predictive relationship between volume difference of those regions and recollection memory performance. Thus, while manual segmentation of the hippocampus is a necessary technique, to researchers or clinicians who do not have access to the needed human expertise its use may be infeasible.

Automated segmentation techniques overcome the need for human expertise by performing segmentations computationally. A popular class of automated methods, *multi-atlas-based segmentation*, rely on a set of expertly labeled neuroanatomical atlases. Each atlas is warped to fit a subject's neuroanatomy using nonlinear registration techniques [3, 11]. Atlas labels are then transformed by this warping and a *label fusion* technique, such as voxel-wise voting, is used to merge the competing label definitions from each atlas into a final segmentation for a subject.

Many descriptions of multi-atlas-based segmentation algorithms report relying on an atlas library containing anywhere between 30 and 80 expertly labeled brains [9, 5, 1, 12, 13]. As noted, the production of an atlas library requires significant manual effort, and is limited since the choice of atlases or segmentation protocol may not reflect the underlying neuroanatomical variability of the population under study or be suited to answer the research questions at hand.

In this paper we propose an automated segmentation method to address the above issues of existing multi-atlas-based methods. Principally, our method aims to dramatically reduce the number of manually labelled atlases needed (under 10). This is achieved by using the small atlas library to boot-strap a much larger "template library", which is then used to segment the subjects in a similar fashion to basic multi-atlas segmentation. This approach has the additional

advantage of using the unique subject population on hand to initialize the segmentation process and improve accuracy.

The essential insight of generating a template library is not new. Heckemann [8] compared generating a template library from a single atlas to standard multi-atlas segmentation and found poor performance and so deemed the approach as inviable. The LEAP algorithm [17] proceeds by iteratively segmenting the unlabelled image most similar to the atlas library images and then incorporating the now-labelled image into the atlas library, but requires 30 starting atlases. The novelty of our method is to demonstrate the possibility of producing comparable segmentation accuracy to these and other multi-atlas-based methods while using significantly fewer manually created atlases.

In our previous work [2], we applied MAGeT brain to segmentation of the human striatum, globus pallidus, and thalamus using a single histologically-derived atlas. The main contribution of this paper is to extend our approach to the human hippocampus and perform a thorough validation over a range of atlas and template library sizes, which was not done in our previous work. Due to the small number of atlases required, our method can easily accommodate different hippocampal definitions. Our aim is not to improve on segmentation accuracy beyond existing methods, but instead to provide a method that trades off manual segmentation expertise for computational processing time while providing sufficient accuracy for clinical and research applications.

2 Methods

MAGeT Brain Algorithm

In this paper, we use the term *atlas* to mean any manually segmented MR image, and the term *atlas library* to mean a set of such images. We use the term *template* to refer to any MR image, and associated labelling, used to segment another image, and the term *template library* to refer to a set of such images. An atlas library may be used as a template library but, as we will discuss, a template library may also be composed of images with computer generated labellings.

The simplest form of multi-atlas segmentation combines labellings derived from several atlases by way of label fusion[5]. We will refer to this as *basic multi-atlas segmentation*. The primary steps are as follows. An atlas library and unlabelled MR images are given as input. Each atlas image is nonlinearly registered to each unlabelled image, and then each atlas' labels are propagated via the resulting transformations. The resulting labels are fused to produce a single, definitive segmentation by some label fusion method (for exampl, by voxel-wise majority voting).

The segmentation approach we propose is best understood as an extension of MAGeT brain adds a preliminary stage to this process in which the template library is constructed rather than given as input. As before, MAGeT brain accepts an atlas library and unlabelled MR images as input. Images for the template library are selected from

To create the template library, labels from each atlas image are propagated to each template library image via the transformation resulting from a non-linear registration between pair of images. As a result, each template library image has a label from each atlas. Basic multi-atlas segmentation is then used to produce segmentations for the entire set of unlabelled images (including those images used in the template library).

Source code can be found at http://github.com/pipitone/MAGeTbrain.

Algorithm 1 Pseudocode for the MAGeT Brain algorithm

```
function BasicMultiAtlasSegmentation(Templates, Subjects)
    for all subject do
        for all template do
            propagate all labels for template to subject space
            store subject labels
        end for
        fuse subject labels
    end for
end function

function MAGeTBrain(Subjects, Atlases, n)
    for i = 1 → n do
        choose a subject to be used as a template
        propagate labels from each atlas to template space
        store the template with all of its labels
    end for
    MultiAtlas(Templates, Subjects)
end function
```

2.2 Subjects

2.2.1 ADNI-1 1.5T Screening

2.2.2 SZ First Episode Patients

Table 1: ADNI-1 1.5T Screening demographics

	N	CN $N = 414$	EMCI $N = 279$	LMCI $N = 560$	AD $N = 278$	Combined $N = 1531$
Age at baseline Years	1531	70.9 74.0 78.4	66.0 71.0 76.2	69.4 74.3 79.4	70.9 75.9 80.9	69.6 74.0 79.2
Sex : Female	1531	50% (206)	44% (123)	38% (215)	45% (125)	44% (669)
Education	1531	14.0 16.0 18.0	14.0 16.0 18.0	14.0 16.0 18.0	12.0 16.0 17.8	14.0 16.0 18.0
Ethnicity : Unknown	1531	1% (4)	1% (2)	1% (4)	1% (3)	1% (13)
Not Hisp/Latino		96% (398)	95% (265)	96% (540)	96% (267)	96% (1470)
Hisp/Latino		3% (12)	4% (12)	3% (16)	3% (8)	3% (48)
CDR-SB	1531	0.0 0.0 0.0	0.5 1.0 1.5	1.0 1.5 2.0	3.5 4.5 5.0	0.0 1.0 2.5
ADAS 13	1511	6.0 9.0 12.0	8.0 12.0 16.0	14.3 18.7 23.0	24.0 29.0 34.0	10.0 15.3 23.0
MMSE	1531	29.0 29.0 30.0	28.0 29.0 30.0	26.0 27.0 29.0	21.2 23.0 25.0	26.0 28.0 29.0

a b c represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables.
N is the number of non–missing values.
Numbers after percents are frequencies.

Table 2: Schizophrenia First Episode Patient Demographics

	N	FEP $N = 81$
Age	80	21 23 26
Gender : M	81	63% (51)
Handedness : ambi	81	6% (5)
left		5% (4)
right		89% (72)
Education	81	11 13 15
SES : lower	81	31% (25)
middle		54% (44)
upper		15% (12)
FSIQ	79	88 102 109

a b c represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables.
N is the number of non–missing values.
Numbers after percents are frequencies.

2.3 Image pre-processing

Before images were registered, the N3 algorithm [16] is first used to minimize the intensity nonuniformity in each of the atlases and unlabeled subject images.

2.4 Registration

2.4.1 Automatic Normalization and Image Matching and Anatomical Labeling (ANIMAL)

The ANIMAL algorithm carries out Image registration is in two phases. In the first, a 12-parameter linear transformation (3 translations, rotations, scales, shears) is estimated between images using an algorithm that maximizes the correlation between blurred MR intensities and gradient magnitude over the whole brain [4]. In the second phase, nonlinear registration is completed using the ANIMAL algorithm [3]: an iterative procedure that estimates a 3D deformation field between two MR images. At first, large deformations are estimated using blurred version of the input data. These larger deformations are then input to subsequent steps where the fit is refined by estimating smaller deformations on data blurred with a Gaussian kernel with a smaller FWHM. The final transformation is a set of local translations defined on a bed of equally spaced nodes that were estimated through the optimization of the correlation coefficient.

For the purposes of this work we used the regularization parameters optimized in Robbins et al. [16].

to MNI website

command line

2.4.2 Automatic Normalization Tools (ANTS)

ANTs is a diffeomorphic registration algorithm which provides great flexibility over the choice of transformation model, objective function, and the consistency of the final transformation. The transformation is estimated in a hierarchical fashion where the MRI data is subsampled, allowing large deformations to be estimated and successively refined at later hierarchical stages (where the data is subsampled to a finer grid). The deformation field and the objective function are regularized with a Gaussian kernel at each level of the hierarchy. The ANTs algorithm is freely available http://www.picsl.upenn.edu/ANTS/. We used an implementation of the ANTS algorithm compatible with the MINC data format, mincANTS https://github.com/vfonov/mincANTS.

We used the following command line when running the ANTS command,

```
mincANTS 3 -m PR[target_file.mnc,source_file.mnc,1,4]
 --number-of-affine-iterations 10000x10000x10000x10000x10000
 --affine-gradient-descent-option 0.5x0.95x1.e-4x1.e-4
 --use-Histogram-Matching --MI-option 32x16000
 -r Gauss[3,0] -t SyN[0.5] -i 100x100x100x20
 -o transformation.xfm
```

These settings were adapted from the "reasonable starting point" given in the ANTS manual.

include a refe manual?

2.5 Label Fusion

Label fusion is term given to the process of combining the information from several candidate labellings for an MR image into a single labelling. In this paper we explore the benefits of three different fusion methods.

2.5.1 Voxel-wise Majority Vote

Labels are propagated from all template library images to a subject. Each output voxel is given the most frequent label at that voxel location amongst all candidate labellings. Ties are broken arbitrarily.

2.5.2 Cross-correlation Weighted Majority Vote

An optimal combination of subjects from the template library has previously been shown to improve segmentation accuracy [1, 5]. In this method, each template library image is ranked in similarity to each unlabelled image by the normalized cross-correlation (CC) of image intensities after linear registration, over a region of interest (ROI) generously encompassing the hippocampus. Only the top ranked template library image labels are used in a voxel-wise majority vote. The ROI is heuristically defined as the extent of all atlas labels after linear registration to the template, dilated by three voxels [2]. The number of top ranked template library image labels is a configurable parameter.

The `xcorr_vol utility` from the ANIMAL toolkit is used to calculate the cross-correlation similarity measure.

2.5.3 Normalised Mutual Information Weighted Majority Vote

This method is similar to cross-correlation weighted voting except that image similarity is calculated by the normalised mutual information score over the region of interest.

is there a refe NMI voting?

The `itk_similarity` utility from the EZMinc toolkit is used to calculate the normalised mutual information measure.

2.6 Goodness-of-fit

Two segmentation can be compared using the Dice Kappa (κ) overlap metric:

$$\kappa = \frac{2a}{2a + b + c}$$

where a is the number of voxels common to both segmentations and $b + c$ is the sum of the voxels uniquely identified by either segmentation. Using this metric we can compare an automatically generated segmentation to a gold-standard segmentation.

2.7 Experiments

Experiments were performed to assess the performance of MAGeT brain with various parameter settings as well as on diverse datasets. In each experiment we contrast the performance of MAGeT brain with that standard single- and multi-atlas segmentations derived from the same atlas library.

2.7.1 ADNI-1 cross-validation

To test the accuracy of the MAGeT brain algorithm with different parameter settings, repeated random sub-sampling cross-validation (RRSCV) was performed on a subset of the ADNI-1 dataset.

Dataset evaluated. 69 1.5T images were randomly selected from the *ADNI1:Screening 1.5T* standardized dataset. Demographics for this subset are shown in Table X .

Atlas and template libary. Atlases consisted of images taken from the dataset, with corresponding manual labels provided by SNT. Atlas library size was varied from 3 to 9 images. The remaining images were segmented, with the template library size varying from 3 to 20 images. Template library images were selected randomly from the images to be segmented.

Registration method. Both the ANTS and ANIMAL registration methods were used.

Label fusion. Majority vote, Cross-correlation weighted majority vote, and Normalized Mutual Information weighted majority vote are used. With the weighted majority vote fusion methods, the number of top labels used in the fusion was varied from 3 to 20 images.

Evaluation. Repeated random sub-sampling cross-validation (RRSCV) consists of repeated trials in which items from the dataset are randomly assigned to a training set or validation set. In each trial, performance on the validation set is measured, and then averaged across all trials.

We performed RRSCV on each combination of parameters listed above: atlas library size, template library size, registration method, and label fusion method. We performed 10 trials per parameter combination. In each validation trial, the training set consisted of the images used as atlases, and the validation set consisted of the images to be segmented. The MAGeT brain algorithm and the basic multi-atlas segmentation procedure were applied to segment the images in the validation set. Additionally, in each trial, the single-atlas segmentation was obtained for each atlas-template.

The gold-standard for the segmentation accuracy of images in the validation set was the SNT manual labels.

2.7.2 ADNI-1 Screening Validation

To test the accuracy of MAGeT brain on a real-world task we segment the entire ADNI-1 dataset using an atlas set that is not representative of the subject set.

Dataset evaluated. All images from the *ADNI1:Screening 1.5T* standardized dataset.

Atlas and template libary. The atlas library consisted of the entire Winterburn atlas set. The Winterburn atlases are digital segmentations of the hippocampus in five in-vivo 300u isotropic T1-weighted MR scans, and include subfield segmentations for the cornus ammonis (CA) 1, CA4, dentate gyrus, subiculum, and CA 2 and 3 combined. Subjects in the Winterburn atlases range in age from 29-57 years (mean age of 37), and include two males and three females.

The template library consisted of 21 randomly selected images from the ADNI1 data dataset (7 healthy, MCI and AD subjects).

Registration method. ANTS, as it performed best in the cross-validation experiment.

Label fusion. Majority vote, as it is simpliest to run and performed equally well in cross-validation experiment.

Evaluation.

Since hippocampal segmentation protocols differ between the ADNI labels and Winterburn atlases, this poses a problem for direct similarity comparisons between labels produced by MAGeT brain and the ADNI labels.

To evaluate the performance of MAGeT brain, we correlate our segmentation volumes with manual segmentation volumes, as well as with hippocampal vol-

ME: table reference

NIMAL, images ed to TAL space

show off the r ber of registra tions/compar did

Is it okay to r results whilst Methods?

explain why v resegment the images with t protocol and directly like t

Table 3: ADNI-1 cross-validation subset demographics

	CN $N = 23$	LMCI $N = 23$	AD $N = 23$	Combined $N = 69$
Age at baseline Years	72.2 75.5 78.5	71.0 77.1 81.4	71.7 77.8 81.8	71.5 76.6 81.3
Sex : Female	43% (10)	43% (10)	43% (10)	43% (30)
Education	16.0 16.0 18.0	15.0 16.0 18.0	12.0 16.0 16.5	14.0 16.0 18.0
Ethnicity : Unknown	0% (0)	0% (0)	0% (0)	0% (0)
Not Hisp/Latino	100% (23)	100% (23)	100% (23)	100% (69)
Hisp/Latino	0% (0)	0% (0)	0% (0)	0% (0)
CDR-SB	0.00 0.00 0.00	0.75 1.50 1.50	4.00 4.50 5.00	0.00 1.50 4.00
ADAS 13	4.67 5.67 12.34	14.34 16.00 20.50	23.83 29.00 31.66	10.00 16.00 25.33
MMSE	28.5 29.0 30.0	25.0 27.0 28.0	21.0 23.0 24.0	24.0 27.0 29.0

a b c represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables.
Numbers after percents are frequencies.

ude a description
lidation process..
that we could use
`validate` or t-test:
rast with QDA
DA (Coupe 2011)
in LOOCV, with
/our segmentations

s of established automated segmentation meth-

dditionally, we compared classification accuracy
ubjects by diagnosis based on hippocampal vol-
using both the SMT labels and our produced
ls.

2.7.3 SZ First Episode Patient Validation

r to the FEP de-
tion above... or
include it here

Dataset evaluated. To validate that MAGeT per-
mance generalises to other diseases, we measure
performance using the best parameter settings
ious found, on a dataset consiting of first episode
zophrenia patients. .

Atlas and template libary. - two different atlas sets: a manual hippocampal segmentation of patients, and Winterburn atlas set.

Registration method. ANTS.

Label fusion. Majority vote.

Evaluation. We validate the FEP-atlas segmentations using Dice's Kappa, and the Winterburn-atlas segmentations by correlating volumes.

2.7.4 Winterburn Atlases Validation

Dataset evaluated. - T1 BRAVO scans of the same subjects included in the Winterburn atlas set. These scans are taken within .... weeks of the scans for the Winterburn atlases.

Atlas and template libary. - Atlas library is Winterburn T1 atlases. Template library consists of all five T1 BRAVOs, plus 15 T1 healthy control images.

Registration method. ANTS

Label fusion. Majority vote.

Evaluation. - Leave one out cross-validation (LOOCV) in which all five subjects are segmented in separate runs of MAGeT brain. In each run, the subject to be segmented is excluded from the Atlas library (so only four atlases are used).

Segmentation accuracy is judged by difference in hippocampal volume.

3 Results

3.1 ADNI-1 Cross-Validation

- find significant improvement over multi-atlas performed with the same parameters. Also, find smoothed performance is monotonically increasing but asymptotic in size of both template and atlas library, with peak performance reached after 15 templates.

Can we statis
capture what
performance i
thing like, the
at which gain
statistically in
cant?

Ideas:

- more atlases -¿ better performance
- larger template library -¿ better performance, but tails off around 10-15 templates
- no significant difference between majority or weighted vote methods (haven???t tested this statistically though).
- consistently performs better than average naive performance by XXX
- using ANTS, with a large enough template library (¿12) MAGeT brain performs better than the average multi-atlas approach with the same number of atlases. using ANIMAL, 5 or more atlases needed before boost seen.
- more atlases -¿ smaller template library required to improve on average multi-atlas performance
- discuss variance? best/worst case? –how often do we expect random template library selection to work decently

show cost (in
tions) / benef
off graph: sho
ber of registra
per Kappa? c
of manual lab
Kappa?)

Table 4: Multi-atlas means

merge	Atlases	ANTS	ANIMAL
1	3.00	0.81	0.76
2	4.00	0.79	0.75
3	5.00	0.82	0.79
4	6.00	0.82	0.78
5	7.00	0.83	0.80
6	8.00	0.83	0.79
7	9.00	0.84	0.80

3.2 ADNI-1 Screen Validation

Ideas:

- A2A shows that if atlas population strongly(?) represents subject set variability, then free choice from atlas population will produce improvements (we know this b/c of extensive validation trials).
- what about in the case where atlas population doesn???t strongly represent subject set variability (e.g. a priori atlas set)? then, we can use atlas selection to refine atlas set?

pa against our
ual rater is low...
ain that

3.3 First Episode Schizophrenic Patients

High volume correlation between Winterburn segmentation volumes and ground truth. (High-ish?) Kappa when using manual segmentations as Atlases.

3.4 Winterburn Atlases Validation

4 Discussion

5 Conclusion

References

[1] P. Aljabar, R Heckemann, a Hammers, J V Hajnal, and D Rueckert. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *NeuroImage*, 46(3):726–38, July 2009.

[2] M. Chakravarty, Patrick Steadman, Matthijs van Eede, Rebecca Calcott, Victoria Gu, Phillip Shaw, Armin Raznahan, Louis Collins, and Jason P Lerch. Performing label-fusion based segmentation using multiple automatically generated tempaltes. *Human Brain Mapping*, 2012.

[3] D. Collins, C. J. Holmes, T. M. Peters, and A. C. Evans. Automatic 3-D model-based neuroanatomical segmentation. *Human Brain Mapping*, 3(3):190–208, October 1995.

[4] D. Collins, P Neelin, T M Peters, and A C Evans. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of computer assisted tomography*, 18(2):192–205, 1994.

[5] D. Collins and Jens C Pruessner. Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion. *NeuroImage*, 52(4):1355–66, October 2010.

[6] E. Geuze, E. Vermetten, and J D Bremner. MR-based in vivo hippocampal volumetrics: 2. Findings in neuropsychiatric disorders. *Molecular Psychiatry*, 10(2):160, September 2004.

[7] A. Hammers, Richard Allom, Matthias J Koepp, Samantha L Free, Ralph Myers, Louis Lemieux, Tejal N Mitchell, David J Brooks, and John S Duncan. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human brain mapping*, 19(4):224–47, August 2003.

[8] R. Heckemann, J V Hajnal, P Aljabar, D Rueckert, and A Hammers. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *NeuroImage*, 46(3):726–38, July 2006.

[9] R. A Heckemann, Shiva Keihaninejad, Paul Aljabar, Katherine R Gray, Casper Nielsen,

Daniel Rueckert, Joseph V Hajnal, and Alexander Hammers. Automatic morphometry in Alzheimer's disease and mild cognitive impairment. *NeuroImage*, 56(4):2024–37, July 2011.

[10] C. Jack, Frederik Barkhof, Matt A Bernstein, Marc Cantillon, Patricia E Cole, Charles Decarli, Bruno Dubois, Simon Duchesne, Nick C Fox, Giovanni B Frisoni, Harald Hampel, Derek L G Hill, Keith Johnson, Jean-François Mangin, Philip Scheltens, Adam J Schwarz, Reisa Sperling, Joyce Suhy, Paul M Thompson, Michael Weiner, and Norman L Foster. Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. *Alzheimer's & dementia*, 7(4):474–485.e4, July 2011.

[11] A. Klein, Jesper Andersson, Babak A Ardekani, John Ashburner, Brian Avants, Ming-Chang Chiang, Gary E Christensen, D Louis Collins, James Gee, Pierre Hellier, Joo Hyun Song, Mark Jenkinson, Claude Lepage, Daniel Rueckert, Paul Thompson, Tom Vercauteren, Roger P Woods, J John Mann, and Ramin V Parsey. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, 46(3):786–802, July 2009.

[12] K. Leung, Josephine Barnes, Gerard R Ridgway, Jonathan W Bartlett, Matthew J Clarkson, Kate Macdonald, Norbert Schuff, Nick C Fox, and Sebastien Ourselin. Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. *NeuroImage*, 51(4):1345–59, July 2010.

[13] J. Lötjönen, Robin Wolz, Juha R Koikkalainen, Lennart Thurfjell, Gunhild Waldemar, Hilkka Soininen, and Daniel Rueckert. Fast and robust multi-atlas segmentation of brain magnetic resonance images. *NeuroImage*, 49(3):2352–65, March 2010.

[14] A. Mouiha and S. Duchesne. Multi-decade hippocampal and amygdala volume analysis: equal variability and limited age effect. *Neuroscience letters*, 499(2):93–8, July 2011.

[15] J. Poppenk and Morris Moscovitch. A Hippocampal Marker of Recollection Memory Ability among Healthy Young Adults: Contributions of Posterior and Anterior Segments. *Neuron*, 72(6):931–937, December 2011.

[16] S. Robbins, Alan C Evans, D Louis Collins, and Sue Whitesides. Tuning and comparing spatial normalization methods. *Medical image analysis*, 8(3):311–23, September 2004.

[17] R. Wolz, Paul Aljabar, Joseph V Hajnal, Alexander Hammers, and Daniel Rueckert. LEAP: learning embeddings for atlas propagation. *NeuroImage*, 49(2):1316–25, January 2010.

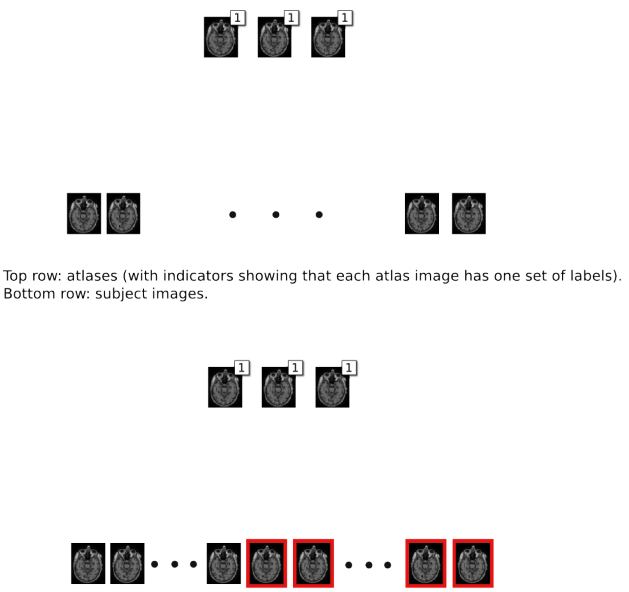

Figure 1: Diagram of the MAGeT Brain algorithm

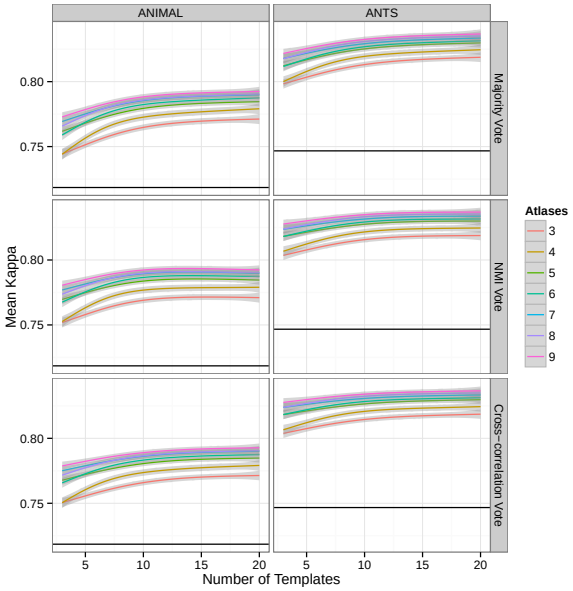

Figure 2: Comparison of MAGeT performance on ADNI-1 subset. Smoothing line fitted using GAM (generalised additive model) from R with defaults from ggplot2 (formula: y s(x, bs = "cs"))

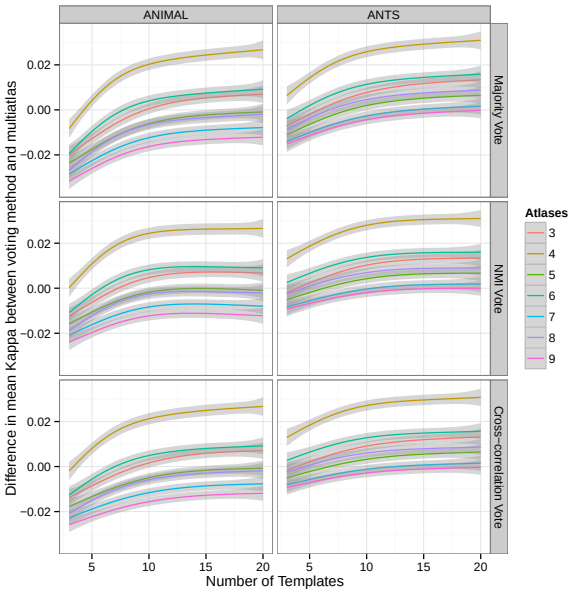

Figure 3: Difference in mean Kappa between MAGeT brain and multi-atlas

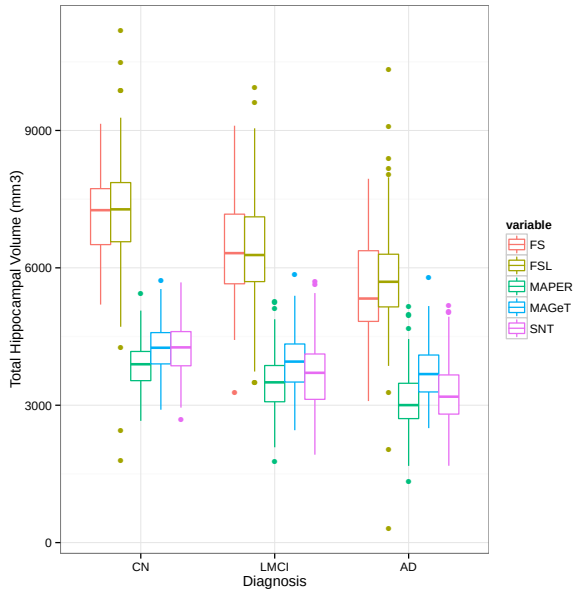

Figure 4: Comparison of HC volumes by FreeSurfer (FSF), MAGeT brain (MAGeT), MAPER, and manual (SNT).

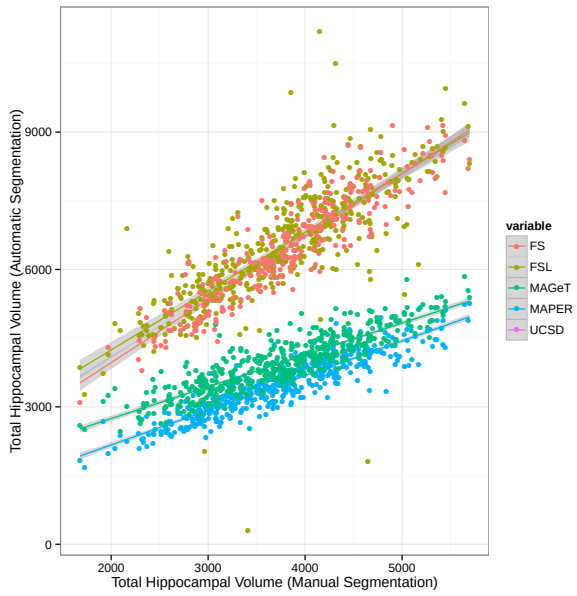

Figure 5: **ADNI Baseline cohort.** Comparison of HC volumes by FreeSurfer (FSF), MAGeT brain (MAGeT), MAPER, and manual (SNT).

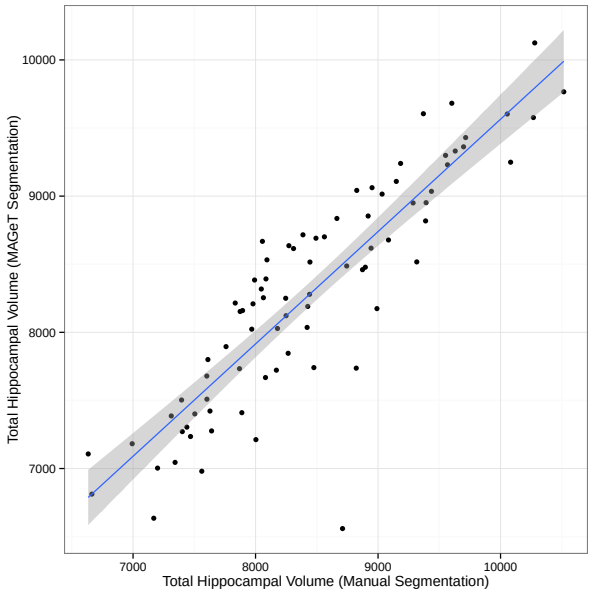

Figure 6: **First Episode Schizophrenic Patients.** Comparison of total HC volumes for MAGeT against manually rated Hippocampal volumes