Segmentation of MRI brain scans using MALP-EM

Christian Ledig[1⋆], Rolf A. Heckemann[2,3], Paul Aljabar[1,4], Robin Wolz[1], Joseph V. Hajnal[4], Alexander Hammers[2,3], Daniel Rueckert[1]

[1] Department of Computing, Imperial College London, London, UK
[2] The Neurodis Foundation, Lyon, France
[3] Centre for Neuroscience (Hammersmith), Imperial College London, UK
[4] Imaging Sciences and Biomedical Engineering, Kings College London, London, UK

Abstract. We employ a modification of our previously published method based on multi-atlas label propagation (MALP) and intensity-based refinement through expectation-maximization (EM) to segment magnetic resonance (MR) brain scans of the OASIS database. We had gold-standard segmentations available for 15 subjects of the same database, which we used as atlases in a multi-atlas propagation setup. After propagating the available atlases using transformations obtained with the robust MAPER approach, we use a locally weighted fusion strategy to merge the 15 atlas label sets into a consensus probabilistic segmentation of the unseen image. We use these probabilistic labels as priors in a subsequent EM refinement step, where we improve the segmentations based on the intensity distribution of the images. On top of the common EM refinement we apply a statistical correction based on the intensity characteristics of each individual region. The intensity profile of certain regions and their individual neighborhoods are not suited for an intensity based EM refinement nor a statistical correction. Therefore, we only refine regions for which intensity based refinement is beneficial and obtain a final segmentation by merging the labels obtained through MALP, MALP-EM and the statistical corrected MALP-EM regions. For evaluation, we segment MR brain scans of 20 subjects of the OASIS database.

1 Introduction

The segmentation of brain images into anatomical regions in magnetic resonance (MR) scans is an important task in neuroimaging. It yields regional volumetric information and labeling of different brain structures which can support clinical decision making. Even though manual annotations by a trained specialist are accurate, they are not scalable, time consuming and thus expensive. A fully automated method that calculates brain segmentations without user interaction is thus highly desirable and the basis for the segmentation of large data sets,

⋆ This work is partially funded under the 7th Framework Programme by the European Commision (http://cordis.europa.eu/ist/).

such as the data from the Alzheimer's Disease Neuroimaging Inititative (ADNI, adni.loni.ucla.edu) [1] or OASIS [2].

In this work we employ a recent segmentation method [3] which combines the advantages of both intensity based methods, e.g. [4], and approaches based on multi-atlas label propagation (MALP), e.g. [5], to segment MR brain images of 20 healthy adult subjects of the OASIS database [2]. Our approach refines subject specific spatial priors obtained through MALP and label fusion [6] in a probabilistic intensity model solved via expectation-maximization (EM) [7]. We furthermore refine certain regions based on statistical intensity characteristics.

2 Method

2.1 Material

We used the dataset provided through the "MICCAI 2012 Grand Challenge and Workshop on Multi-Atlas Labeling". The training dataset consists of 15 T_1-weighted images with corresponding labels created by experts[1]. We segmented a testing dataset consisting of 20 otherwise identical T1-weighted images with hidden labels into 138 regions. The performance of our approach was evaluated using an automatic online evaluation interface provided through the Challenge.

2.2 Multi-Atlas Label Propagation with EM refinement (MALP-EM)

We use multi-atlas label propagation to derive a subject-specific probabilistic brain atlas for an unseen T_1 weighted MR scan $\mathbf{I}$ that is to be segmented. We incorporate these probabilistic labels into our EM framework as spatial anatomical priors. We index the n voxels of $\mathbf{I}$ by $i = 1, \ldots, n$, so that for intensities $y_i \in \mathbb{R}$ an image can be defined as $\mathbf{I} = \{y_1, y_2, \ldots, y_n\}$. The probabilistic priors are created by transforming M manually generated atlases to the coordinate space of the unseen image. We calculate the M transformations for the label propagation with a non-rigid registration method based on free-form deformations (FFD) [8, 9], which follows a preceding rigid and affine alignment. In particular we employed MAPER [10], which incorporates tissue probability maps into the registration. The probabilistic atlas is then created with a locally weighted multi-atlas fusion strategy [6], by employing a Gaussian weighted sum of squared differences on rescaled, intensity-normalized images. We followed the approach of van Leemput et al. [7] and estimated the hidden segmentation by means of the observed intensities $\mathbf{y}$. Assuming that the observed log-transformed intensities of voxels belonging to a certain class k are normally distributed with mean μ_k and standard deviation σ_k, yields the model parameters $\mathbf{\Phi} = \{(\mu_1, \sigma_1), (\mu_2, \sigma_2), \ldots, (\mu_K, \sigma_K)\}$. We applied regularization of the resulting segmentation using the approach of global and stationary Markov Random Fields (MRF) described in [11].

[1] provided by Neuromorphometrics, Inc. (http://Neuromorphometrics.com/) under academic subscription.

2.3 Statistical correction of MALP-EM

During our experiments we observed that the EM algorithm tends to produce segmentations with a too low intensity variance within a region (intra-class variance) compared to the gold-standard segmentations. We therefore calculated an expected normalized intra-class variance for each region ($\sigma^2_{\mathrm{Gold},k}$) by averaging the normalized standard deviations $\frac{\sigma_k}{\mu k}$ of each class over the training subjects. We furthermore calculated the averaged (average over all training subjects segmented with a leave-one-out strategy) normalized standard deviation within each region produced by the EM algorithm ($\sigma_{\mathrm{EM},k}$) . By calculating $\Delta_k = (\sigma_{\mathrm{Gold},k} - \sigma_{\mathrm{EM},k})^2$ we estimated by which value the intra-class variance of a certain class should be increased in average to better match the gold-standard characteristics. In a subsequent refinement step we then corrected the intra-class statistics of each class by adding voxels with posterior probability greater than 10%, in decreasing order regarding the label probability, to the region unless the intra-class variance increased by Δ_k. Overlaps of most cortical regions with the gold-standard could be improved using statistical correction.

2.4 Fusion of MALP and MALP-EM

Our experiments revealed that some regions are ill-suited for intensity based refinement, due to either their intensity properties, or to those of their neighborhood. For example, no improvements using EM were obtained for the structures thalamus and putamen which can be explained with the wide overlap of their intensity profile with the profile of white matter. This is also shown in [3]. For these structures it is preferable to rely on the segmentation obtained through MALP alone. By segmenting all available training datasets with a leave-one-out strategy, we determined the subset of regions for which the standard EM refinement or the statistically corrected version is beneficial. We then created a final segmentation by combining the refined labels for this subset with the labels from the MALP approach for the remaining regions. In case of overlapping regions, we labeled a voxel according to the EM-refined label.

2.5 Parameters

To identify neighbouring tissue classes for the implementation of the MRF, we counted the labels of adjacent voxels in the gold-standard segmentations. After thresholding we obtained a 139×139 adjacency matrix G that describes the MRF, with entry (i, i) eqauls 0 and entry (i, j) defined as 1.0 if structures i and j share a boundary and 1.5 if structures i and j are distant. For a voxel size of 1x1x1mm we set for the locally weighted fusion the parameter σ to 2.5. Parameters were optimized using a leave-one-out strategy on the training datasets.

3 Results

The presented approach was evaluated using 20 datasets of the OASIS database with hidden labels. The results were automatically calculated through the Grand

Challenge on Multi-Atlas Labeling. We observe that the MALP approach performs very well on most of the 36 subcortical regions (average Dice similarity coefficient greater than 85%). Since the cohort consists of healthy adults with little intersubject variability, it is not surprising that registration based approaches perform well on this dataset. The EM- and statistical-based refinement is thus particularly relevant in cortical regions where, due to the high structural variability within the brain, registration based approaches are less accurate. Also the high intensity contrast at the cortical boundary between white and grey matter tissue is predestined for intensity based EM refinement. We obtain an average Dice coefficient of 73.28% for cortical and 82.52% for subcortical regions on the testing dataset. This yields an overall average label overlap of 75.76%.

References

1. S. G. Mueller, M. W. Weiner, L. J. Thal, et al., "The Alzheimer's Disease Neuroimaging Initiative," *Neuroimaging Clinics of North America*, vol. 15, no. 4, pp. 869–877, 2005.
2. D. S. Marcus, T. H. Wang, J. Parker, et al., "Open access series of imaging studies (oasis): Cross-sectional mri data in young, middle aged, nondemented, and demented older adults," *Journal of Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498–1507, 2007.
3. C. Ledig, R. Wolz, P. Aljabar, et al., "Multi-class brain segmentation using atlas propagation and EM-based refinement," *Proc. of ISBI 2012*, pp. 896–899, 2012.
4. J. MP. Lötjönen, R. Wolz, J. R. Koikkalainen, et al., "Fast and robust multi-atlas segmentation of brain magnetic resonance images," *NeuroImage*, vol. 49, no. 3, pp. 2352–2365, 2010.
5. R. A. Heckemann, S. Keihaninejad, P. Aljabar, et al., "Automatic morphometry in Alzheimer's disease and mild cognitive impairment," *NeuroImage*, vol. 56, no. 4, pp. 2024–2037, 2011.
6. X. Artaechevarria, A. Munoz Barrutia, and C. Ortiz de Solorzano, "Combination strategies in multi-atlas image segmentation: Application to brain MR data," *IEEE TMI*, vol. 28, no. 8, pp. 1266–1277, 2009.
7. K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based tissue classification of MR images of the brain," *IEEE TMI*, vol. 18, no. 10, pp. 897–908, 1999.
8. D. Rueckert, L. I. Sonoda, C. Hayes, et al., "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE TMI*, vol. 18, no. 8, pp. 712–721, 1999.
9. M. Modat, G. R. Ridgway, Z. A. Taylor, et al., "Fast free-form deformation using graphics processing units," *Computer Methods and Programs in Biomedicine*, vol. 98, no. 3, pp. 278–284, 2010.
10. R. A. Heckemann, S. Keihaninejad, P. Aljabar, et al., "Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation," *NeuroImage*, vol. 51, no. 1, pp. 221–227, 2010.
11. M. J. Cardoso, M. J. Clarkson, G. R. Ridgway, et al., "Load: A locally adaptive cortical segmentation algorithm," *NeuroImage*, vol. 56, no. 3, pp. 1386–1397, 2011.