Published in final edited form as:

Patch Based Tech Med Imaging (2017). 2017 September ; 10530: 138–145. doi: 10.1007/978-3-319-67434-6_16.

Multiple Sclerosis Lesion Segmentation Using Joint Label Fusion

Mengjin Dong¹, Ipek Oguz¹, Nagesh Subbana¹, Peter Calabresi², Russell T. Shinohara³, and Paul Yushkevich¹

¹Penn Image Computing and Science Lab, University of Pennsylvania, Philadelphia, PA, USA

²The Johns Hopkins Calabresi Lab, Johns Hopkins University, Baltimore, MD, USA

³Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, PA, USA

Abstract

This paper adapts the joint label fusion (JLF) multi-atlas image segmentation algorithm to the problem of multiple sclerosis (MS) lesion segmentation in multi-modal MRI. Conventionally, JLF requires a set of atlas images to be co-registered to the target image using deformable registration. However, given the variable spatial distribution of lesions in the brain, whole-brain deformable registration is unlikely to line up lesions between atlases and the target image. As a solution, we propose to first pre-segment the target image using an intensity regression based technique, yielding a set of “candidate” lesions. Each “candidate” lesion is then matched to a set of similar lesions in the atlas based on location and size; and deformable registration and JLF are applied at the level of the “candidate” lesion. The approach is evaluated on a dataset of 74 subjects with MS and shown to improve Dice similarity coefficient with reference manual segmentation by 12% over intensity regression technique.

1 Introduction

Multiple sclerosis (MS) is a degenerative disease of the central nervous system in which the axonal myelin sheath is damaged due to an autoimmune reaction. There is no cure for MS; however, early detection and tracking of the lesions play a vital part in the clinical management of the disease. The burden of the disease (volume of lesions) and the number of lesions in the brain are important metrics in tracking disease and the efficacy of therap.

Manual segmentation of MS lesions is costly and suffers from low reliability due to differences between experts [6]. Consequently, automatic segmentation of lesions using multiple MRI modalities is the preferred approach to quantifying lesion burden. Automatic segmentation of lesions is complicated by variability in the shape, size and position of MS lesions and the differences in the apparent extent of lesions across different MRI modalities. Many techniques for automatic MS lesion segmentation have been proposed and evaluated, as reviewed in [8]. The largest class of techniques performs statistical inference on image

intensity using unsupervised or supervised approaches, and frequently incorporating spatial priors such as Markov random fields; examples include [4,6,11–14].

Recently, patch-based multi-atlas segmentation techniques have been applied to MS lesion segmentation with considerable success [5,9,10]. These methods use large libraries of patches extracted from expert-annotated atlases to match candidate patches in the target image. The advantage of this approach is that it incorporates anatomical context into the segmentation decision: the decision to assign a voxel a label of lesion or non-lesion is based not only on voxel intensity characteristics but on the intensity and texture of the surrounding region. However, a limitation is that they are restrictive in terms of the transformations that atlas patches may undergo when matching the target image. For example, the atlas and the target image may contain a periventricular lesion of similar appearance, but if the ventricle is thin in the atlas and enlarged in the target, the patch similarity metric may be low. Allowing atlas patches to deform, as is proposed in this paper, may improve the segmentation of the target lesion.

Multi-atlas label fusion (MALF) is an approach that a set of expert-labeled atlas images are matched to the target image – typically using non-linear registration – and segmentations from deformed atlas images are fused into a consensus segmentation of the target image [7]. Effective label fusion techniques weigh atlases based on local intensity similarity to the target image [7]. In particular, joint label fusion (JLF) assigns spatially varying atlas weights in a way that accounts for error correlations between pairs of atlases, achieving excellent performance in brain structure segmentation [15]. However, conventional MALF is ill-suited for MS lesion segmentation since whole-brain registration is unlikely to match MS lesions between individuals due to variability in lesion location and extent. In this paper, we adapt JLF to MS lesion segmentation by performing non-linear registration on a regional, rather than whole-brain level. The approach consists of three components:

1. Pre-segmentation of the target image into “candidate” lesions using the OASIS intensity-based technique [13]. For each “candidate” lesion, a region of interest (ROI) is extracted from the MRI. Very large candidate lesions are broken up into smaller overlapping ROIs.
2. Selection of a subset of anatomically similar atlas ROIs for each target “candidate” lesion ROI.
3. Non-linear registration of atlas ROIs to the target ROI and JLF-based segmentation of the target ROI.

These components are detailed below and an evaluation of the proposed approach is carried out in a set of 74 multicontrast MRI scans of patients with MS. The results indicate that regional JLF segmentation improves accuracy relative to manual segmentation over intensity-based segmentation alone. To our knowledge, this is the first application of MALF with deformable registration to the problem of MS lesion segmentation.

2 Materials and Methods

2.1 Imaging Data

Analysis is carried out in a cross-sectional dataset of 3 tesla T1-weighted and FLAIR MRI acquired in 94 MS patients. The generation of OASIS also uses T2-weighted and PD-weighted. MS lesions were segmented manually in all images by a trained expert with 10+ years of experience. MRI volumes were skull stripped via the algorithm SPECTRE and then bias field corrected using the N3 algorithm. FLAIR volumes were rigidly aligned to the T1-weighted MRI. Figure 1 shows an example of the image data.

2.2 Lesion Pre-segmentation

The goal of pre-segmentation is to produce a pool of “candidate” lesions that includes most of the “true” lesions in the target image but may also include false positives. In subsequent steps, MALF is used to refine this segmentation, with the goal of refining lesion boundaries and removing false positives. Pre-segmentation is produced using the OASIS approach, which fits a logistic regression model to the normalized T1 and FLAIR MRI intensity of lesion and non-lesion voxels in the training set [13]. Thresholding this probability map at a given level ρ gives a binary segmentation of the lesions. Sweeney et al. report the sensitivity and specificity of OASIS for different levels of ρ [13]. OASIS was trained on 20 of the 94 subjects’ data and applied to the remaining 74 subjects. Based on previous experience with the OASIS method and visual inspection of the data in the training set, we adopted a threshold of $\rho = 0.25$. The 20-subject OASIS training data was not used in subsequent experiments.

Candidate lesions are obtained by splitting the binary OASIS output into connected components. Components smaller than 3 voxels were discarded. Additionally, we broke up large components into separate sub-components, because in some subjects periventricular lesions yield a single connected component and in others they yield several connected components, and the matching strategy discussed below is based on global morphological features (size, location) that match up poorly between contiguous and split candidate periventricular lesions. To split connected components we use the SLIC supervoxel algorithm [1] with the approximate supervoxel size of $\frac{1}{10}$ of the image dimensions. Each candidate lesion (or sub-lesion) was then padded by 10 voxels on each side, yielding a set of rectilinear ROIs. Corresponding ROIs were extracted from the FLAIR and T1-weighted images. We refer to these as “lesion ROIs”.

2.3 MALF Approach and Cross-Validation Overview

The dataset of 74 subjects processed by OASIS was randomly split into five training/testing folds with 59 training image and 15 test images (the fifth fold had 60 training and 14 testing images). All subsequent analysis was repeated independently for each fold. Each candidate lesion ROI in each training image of each fold was processed in parallel. For each such “target” lesion ROI, we first select a set of 50 suitable lesion ROIs from the training subset (Atlas Selection) and then perform MALF segmentation using these 50 lesion ROIs as atlases. These steps are described below.

2.4 Atlas Selection

Atlas selection is necessary because the total number of lesion ROIs in the training subset is too large (order of 1000) and most lesion ROIs are contextually different from a given target lesion. To perform atlas selection, we compare each OASIS-extracted candidate lesion in the training subset to target candidate lesion based on three factors: (1) lesion size (after splitting by SLIC algorithm); (2) distance from the lesion center to the lateral ventricles; (3) location of the closest point on the lateral ventricles. To compute (2) and (3), we register each subject's T1-weighted MRI to a whole-brain template, and warp a surface mesh representation of the ventricles defined in template space into the subject's space. For each lesion i , we then find a vertex on the ventricle surface $\mathbf{x}_v^i$ that is closest to the lesion center $\mathbf{x}_c^i$. Given two lesions i and j , we compute the triple of similarity metrics,

$$\left(\left| \log \left(\frac{V_i}{V_j} \right) \right|, \left| \frac{|\mathbf{x}_v^i - \mathbf{x}_c^i|}{\sqrt[3]{V_i^{\text{TB}} - V_i^{\text{LV}}}} - \frac{|\mathbf{x}_v^j - \mathbf{x}_c^j|}{\sqrt[3]{V_j^{\text{TB}} - V_j^{\text{LV}}}} \right|, d_v(\mathbf{x}_v^i, \mathbf{x}_v^j) \right),$$

where V_i is the volume of the lesion i , V_i^{TB} is the total brain volume, V_i^{LV} is the lateral ventricle volume, and $d_v(\mathbf{x}_v^i, \mathbf{x}_v^j)$ is the geodesic distance between ventricle mesh vertices corresponding to $\mathbf{x}_v^i$ and $\mathbf{x}_v^j$ in template space. Lesions are considered contextually similar if the three metrics are low. Since the metrics have different units, for a given target lesion i , we rank all training lesions separately on each metric, resulting in a triple of rank values. The 50 training lesions with the smallest mean rank value are “atlases” for segmenting the target lesion.

The above atlas selection strategy is motivated by the importance of the ventricles in providing context for MS lesion segmentation. Many MS lesions are adjacent to the ventricles (periventricular) and the shape, size and appearance is likely to be similar at the same location relative to the ventricles. The ranking strategy may select atlases that not very similar to the target ROI, but the JLF method will weight less on them.

2.5 Joint Label Fusion

Given a target lesion ROI and a set of 50 training lesion ROIs obtained by atlas selection, the standard JLF approach is employed. Each atlas lesion ROI is registered to the target ROI using a greedy diffeomorphic registration method [3]. Registration is initialized by matching the centers of mass of the candidate lesions between the atlas and target ROIs. Deformable registration maximizes the local cross-correlation metric with a $9 \times 9 \times 9$ voxel window and is regularized by Gaussian smoothing of the deformation field at each iteration, as in [2]. The T1-weighted and FLAIR images contribute equally to the metric computation.

After registration, the atlas FLAIR and T1-weighted ROIs are resliced into the target image space. The corresponding expert segmentation is also warped to target image space (although some atlas ROIs may be false positives and contain empty expert segmentations).

The warped images and segmentations are input to the JLF algorithm [15], which produces a spatially varying voting weight map $w_i(\mathbf{x})$ for each atlas i . To compute the weights $\{w_i\}$ at a voxel $\mathbf{x}$, the total expected segmentation error is minimized with respect to $\{w_i\}$. The total segmentation error is expressed as pairwise expected joint segmentation errors for all pairs of atlases $\{i, j\}$. The latter is approximated by considering patch similarities between atlases i and j and the target image, as described in [15].

3 Results

Segmentation was run on a multi-CPU cluster with 30 parallel processes each using two cores. For 15-image test set, segmentation required about an hour, with most time spent in deformable registration and JLF computation.

Figure 2 shows an example of the segmentation by the proposed method, compared with segmentation by OASIS alone and expert manual segmentation. Table 1 reports, for each cross-validation fold, the mean Dice similarity coefficient between the expert segmentation and the automatic segmentations obtained by OASIS alone and the proposed method, with OASIS threshold $\rho = 0.25$. The mean Dice coefficient across all five folds was 0.52 ± 0.22 for OASIS and 0.59 ± 0.19 for the proposed method, corresponding to roughly 12% improvement. The 95% confidence interval on the improvement in Dice coefficient is (0.04, 0.09), i.e., the improvement is highly significant.

We also observed that segmentation performance varies by the total lesion burden in each subject (i.e., total volume of all expert-segmented lesions). Figure 3 shows a scatter plot of OASIS vs. JLF Dice coefficient for different testing subjects, with the total lesion burden indicated for each subject. Overall, Dice coefficient is higher in subjects with greater lesion burden. The improvement from JLF is more notable in subjects with lesser lesion burden, suggesting that JLF improves the segmentation of small lesions more than large lesions.

Modulating the OASIS threshold ρ can increase the number and size of candidate lesions. Values of ρ below 0.25 should result in over-segmentation of the lesions, on average; and it is possible that by removing false positives, JLF would perform better for smaller values of ρ than the optimal OASIS threshold. Figure 4 plots the Dice coefficient for OASIS and JLF for different values of ρ between 0.2 and 0.3. However, surprisingly, there is no apparent advantage to using smaller threshold values for JLF accuracy.

4 Conclusion

We proposed an approach for MS lesion segmentation that behaves as a wrapper around an intensity-based lesion segmentation technique (OASIS) and performs multi-atlas segmentation separately for each candidate lesion ROI identified by the intensity-based technique. The approach is the first application of the joint label fusion algorithm to MS lesion segmentation. The addition of the JLF improves MS lesion segmentation performance by 12% over OASIS, which is considered a state of the art technique. Future work will focus on using other approaches to generate candidate lesions, comparisons on grand challenge data, and extensions to 4D lesion modeling.

Acknowledgments

This work was supported, in part, by the NIH grants NIBIB R01EB017255, NINDS R01NS082347, NINDS R01NS085211, NINDS R21NS093349, NINDS R01NS094456.

References

1. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. SLIC super-pixels compared to state-of-the-art superpixel methods. *IEEE Trans Pattern Anal Mach Intell.* 2012; 34(11):2274–2282. [PubMed: 22641706]
2. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal.* 2008; 12(1):26–41. [PubMed: 17659998]
3. Avants B, Gee JC. Geodesic estimation for large deformation anatomical shape averaging and interpolation. *Neuroimage.* 2004; 23(Suppl 1):S139–S150. [PubMed: 15501083]
4. Freifeld O, Greenspan H, Goldberger J. Multiple sclerosis lesion detection using constrained GMM and curve evolution. *Int J Biomed Imaging.* 2009; 2009:715124. [PubMed: 19756161]
5. Guizard N, Coupé P, Fonov VS, Manjón JV, Arnold DL, Collins DL. Rotation-invariant multi-contrast non-local means for MS lesion segmentation. *Neuroimage Clin.* 2015; 8:376–389. [PubMed: 26106563]
6. Harmouche R, Subbanna NK, Collins DL, Arnold DL, Arbel T. Probabilistic multiple sclerosis lesion classification based on modeling regional intensity variability and local neighborhood information. *IEEE Trans Biomed Eng.* 2015; 62(5):1281–1292. [PubMed: 25546852]
7. Iglesias JE, Sabuncu MR. Multi-atlas segmentation of biomedical images: a survey. *Med Image Anal.* 2015; 24(1):205–219. [PubMed: 26201875]
8. Lladó X, Oliver A, Cabezas M, Freixenet J, Vilanova JC, Quiles A, Valls L, Ramió-Torrentà L, Rovira A. Segmentation of multiple sclerosis lesions in brain MRI: a review of automated approaches. *Inf Sci.* 2012; 186(1):164–185.
9. Mechrez R, Goldberger J, Greenspan H. Patch-based segmentation with spatial consistency: application to ms lesions in brain MRI. *J Biomed Imaging.* 2016; 2016:3.
10. Roy, S., He, Q., Carass, A., Jog, A., Cuzzocreo, JL., Reich, DS., Prince, J., Pham, D. SPIE Medical Imaging. International Society for Optics and Photonics; 2014. Example based lesion segmentation; p. 90341Y
11. Schmidt P, Gaser C, Arsic M, Buck D, Förschler A, Berthele A, Hoshi M, Ilg R, Schmid VJ, Zimmer C, Hemmer B, Mühlau M. An automated tool for detection of flair-hyperintense white-matter lesions in multiple sclerosis. *Neuroimage.* 2012; 59(4):3774–3783. [PubMed: 22119648]
12. Shiee N, Bazin PL, Ozturk A, Reich DS, Calabresi PA, Pham DL. A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. *Neuroimage.* 2010; 49(2): 1524–1535. [PubMed: 19766196]
13. Sweeney EM, Shinohara RT, Shiee N, Mateen FJ, Chudgar AA, Cuzzocreo JL, Calabresi PA, Pham DL, Reich DS, Crainiceanu CM. OASIS is automated statistical inference for segmentation, with applications to multiple sclerosis lesion segmentation in MRI. *NeuroImage: Clin.* 2013; 2:402–413. [PubMed: 24179794]
14. Van Leemput K, Maes F, Vandermeulen D, Colchester A, Suetens P. Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE Trans Med Imaging.* 2001; 20(8):677–688. [PubMed: 11513020]
15. Wang H, Suh JW, Das SR, Pluta JB, Craige C, Yushkevich PA. Multi-atlas segmentation with joint label fusion. *IEEE Trans Pattern Anal Mach Intell.* 2013; 35(3):611–623. [PubMed: 22732662]

Fig. 1. T1-weighted and FLAIR MRI (axial plane) for one subject, with expert segmentation superimposed in the bottom row.

Fig. 2.
An example dataset with segmentations produced by OASIS, the proposed method (OASIS +JLF) and manually.

Fig. 3.

Comparison of JLF vs OASIS segmentation accuracy by total lesion burden.

Fig. 4. Comparison of Dice similarity coefficients for OASIS and JLF methods using different OASIS thresholds ρ . Average Dice coefficient across all five folds is plotted.

Table 1

The Dice similarity coefficients of the five-fold cross-validation for JLF against the expert segmentations and the corresponding OASIS values on each fold with OASIS threshold $\rho = 0.25$.

Technique	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average
OASIS	0.5160	0.5479	0.5147	0.4855	0.5467	0.5218
JLF	0.5827	0.6221	0.5770	0.5433	0.6062	0.5860