AUTOMATIC LONGITUDINAL MULTIPLE SCLEROSIS LESION SEGMENTATION: MSmetrix

Saurabh Jain[1], *Diana M. Sima*[1], *Dirk Smeets*[1]

[1]ico**metrix**, R&D, Leuven, Belgium

ABSTRACT

Accurate and consistent multiple sclerosis (MS) brain lesion segmentation and volumetry could be an added value to MS clinicians. In this paper, MS**metrix** is presented, an automatic and reliable method, which uses 3D T1-weighted and FLAIR MR images in a probabilistic model to detect white matter lesions as an outlier with respect to the normal brain, while segmenting the brain tissue into grey matter, white matter and cerebrospinal fluid. The actual lesion segmentation is performed based on prior knowledge about the location (within white matter) and the appearance (hyperintense on FLAIR) of lesions. The randomness in longitudinal lesion segmentation for each subject is reduced by harmonising the trade-off between temporal consistency across time points and segmentation diversity at each time point. The method has been validated on the dataset available from the longitudinal MS lesion segmentation challenge 2015.

Index Terms— multiple sclerosis, Magnetic Resonance Imaging, white matter lesions, brain segmentation, MS**metrix**

1. INTRODUCTION

Visual assessment of multiple sclerosis (MS) brain lesions in longitudinal MRI is an important clinical procedure for disease prognosis. Manual delineation of lesions, although most trustworthy, is time consuming, costly and suffers from intra- and inter rater variability. To improve the accuracy of manual segmentation in a follow-up scan, image subtraction between the reference scan and follow-up scan [1] is often used to find new lesions. This is followed by a visual evaluation which is always subjective. Several automatic MS lesion segmentation approaches have been proposed in the past two decades to address the problem of accurate lesion segmentation and its temporal consistency. The first category of methods are single MRI time point based (for example, [2]) that can find lesions by processing each time point separately. The segmentation performance of an automatic method is usually dependent on the image quality, which often differs between the time points in a clinical environment due to different scanners, artefacts etc. This difference in the image quality introduces randomness in the segmentation at each time point, which decreases the consistency of automatic methods. Therefore, the second category of methods take the patient's MRIs from multiple time points and find MS lesions at each time point such that the randomness in the longitudinal segmentation could partially be reduced (for example, [3]). Although our method primarily belongs to the first category, segmentation randomness is reduced by encouraging temporal consistency of spatial neighbours, while still maintaining segmentation diversity (e.g., a lesion appearing or disappearing in-between the time points).

This study has been supported by TRANSACT (FP7-PEOPLE-2012-ITN-316679), CENTER-TBI (FP7-COOPERATION-2013-602150) and BRAINPATH (FP7-PEOPLE-2013-IAPP-612360)

2. METHOD

Fig. 1 shows the architecture of MS**metrix** for the longitudinal MS lesion segmentation. The longitudinal segmentation is performed in two steps: lesion segmentation for each time point individually and temporal consistency. We describe each of the steps into more details below.

2.1. Lesion segmentation

The lesion segmentation [4] takes as inputs 3D T1-weighted image and a pre-processed 3D FLAIR (co-registered to T1) image available from an MS patient. The lesion segmentation step has four stages: brain segmentation, outlier estimation, pruning and lesion filling. In the brain segmentation stage, a probabilistic model is formulated to segment the T1-weighted image into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) using an expectation-maximisation algorithm. In the outlier estimation step, an outlier class is estimated from the co-registered FLAIR image of the same patient using the three tissue class segmentations from the previous step as prior information. This is performed using the same EM algorithm as mentioned above, but now an outlier map is included. This map is iteratively updated by the EM algorithm and, after convergence, an outlier belief image is produced. In the pruning stage, we segment the lesions in the outlier map, i.e., we 'prune' the outlier map, as not every outlier is a lesion. In order to differentiate the MS lesions from such non-lesion outliers, some extra a priori information about the location and the appearance of the lesions needs to

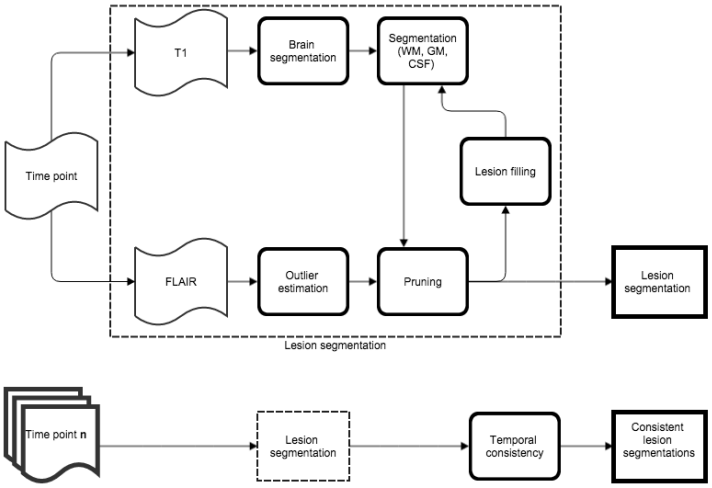

Fig. 1. MS**metrix** architecture for the longitudinal MS lesion segmentation.

be incorporated. The outliers need to be in the WM region and the underlying intensities of the outliers should be hyperintense compared to the GM intensities on bias field corrected FLAIR image. In the lesion filling stage, this lesion segmentation is then used to fill in the lesions in the bias corrected T1-weighted image with WM intensities. These four stages are repeated until convergence and the lesion segmentation is produced as an output.

2.2. Temporal consistency

Since each time point was processed independently, the segmentations could be inconsistent in time. Temporal consistency has two aspects. Firstly, the randomness in temporal segmentation because of varying image quality should be either absent or minimal. Secondly, if a lesion is present on a few or on every time point, it is expected that any automatic method could segment it at each time point. Therefore, the temporal consistency for a voxel p at time point t ($C_{p,t} \in [0,1]$) is defined based on its temporal neighbourhood $N_{p,t}^{time} \in \{t-1, t, t+1\}$, similar to [5] as:

$$C_{p,t} = 1 - \frac{\delta_{N_{p,t}^{time}}}{|N_{p,t}^{time}| - 1} \tag{1}$$

where $\delta_{N_{p,t}^{time}}$ is defined as number of times the segmentation label changes in $N_{p,t}^{time}$ for voxel p. The quantity $\delta_{N_{p,t}^{time}}$ is a measure of randomness in the temporal segmentation. The higher the number, the lower its temporal consistency.

The new temporal consistent label for the voxel p at time point t ($f_{p,t}$) is defined based on the temporal consistency of its $3 \times 3 \times 3$ spatial neighbourhood ($N_{p,t}^{space}$) as follows:

$$f_{p,t} = \begin{cases} f_{p,t} & \text{if } C_p^{avg} \geq 0.5 \\ \text{mode}\{f_Q\} & \text{otherwise} \end{cases} \tag{2}$$

where C_p^{avg} is the average consistency for voxel p based on all time points and $Q = \underset{i \in N_{p,t}^{space}}{\text{argmax}}\, C_{i,t}$. From equation 2, we notice that the $f_{p,t}$ is changed only if C_p^{avg} is < 0.5. This way the diversity at a particular time point is ensured as the voxel is quite consistent in time. If the segmentation label $f_{p,t}$ is less consistent, then it is replaced with the modal value of segmentation labels of its most consistent neighbours.

Encouraging temporal consistency on the challenge's training dataset, our average consistency over the lesions was increased from 65% (range: 59%-68%) to 79% (range: 73%-86%). For comparison, the average consistency for rater-1 for training data was 72% (range: 67%-79%) and for the rater-2 was 67% (range: 58%-75%).

3. REFERENCES

[1] Liguori M. et al., "One year activity on subtraction MRI predicts subsequent 4 year activity and progression in multiple sclerosis," *J. Neurology., Neurosurgury. Psychaitry*, vol. 82, no. 10, pp. 1125–1131, 2011.

[2] Van Leemput K. et al., "Automated segmentation of multiple sclerosis lesions by model outlier detection," *IEEE Transactions on Medical Imaging*, vol. 20, no. 8, pp. 677–688, 2001.

[3] Bosc M. et al., "Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution," 2003.

[4] Jain S. et al., "Automatic segmentation and volumetry of multiple sclerosis brain lesions from MR images," NeuroImage: Clinical (Submitted), 2015.

[5] Zhong Xue et al., "CLASSIC: Consistent longitudinal alignment and segmentation for serial image computing," *NeuroImage*, vol. 30, no. 2, pp. 388 – 399, 2006.