HAL
open science

Automatic Multiple Sclerosis lesion segmentation from Intensity-Normalized multi-channel MRI

Jeremy Beaumont, Olivier Commowick, Christian Barillot

To cite this version:

Jeremy Beaumont, Olivier Commowick, Christian Barillot. Automatic Multiple Sclerosis lesion segmentation from Intensity-Normalized multi-channel MRI. Proceedings of the 1st MICCAI Challenge on Multiple Sclerosis Lesions Segmentation Challenge Using a Data Management and Processing Infrastructure - MICCAI-MSSEG, Oct 2016, Athens, Greece. inserm-01424802

HAL Id: inserm-01424802

https://inserm.hal.science/inserm-01424802v1

Submitted on 2 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Automatic Multiple Sclerosis lesion segmentation from Intensity-Normalized multi-channel MRI

Jeremy Beaumont Olivier Commowick Christian Barillot

VisAGeS U746 INSERM / INRIA, IRISA UMR CNRS 6074, Rennes, France

Abstract. In the context of the FLI MICCAI 2016 MSSEG challenge for lesion segmentation, we present a fully automated algorithm for Multiple Sclerosis (MS) lesion segmentation. Our method is composed of three main steps. First, the MS patient images are registered and intensity normalized. Then, the lesion segmentation is done using a voxel-wise comparison of multi-channel Magnetic Resonance Images (MRI) against a set of controls. Finally, the segmentation is refined by applying several lesion appearance rules.

Keywords: Multiple Sclerosis, Intensity Normalization, Statistics, MRI

1 Introduction

Multiple Sclerosis (MS) is a auto-immune brain degenerative disease causing irreversible patient handicap and which is still not well understood. Lesion detection is a major step to evaluate the patient disease status and its future evolution. Manual and semi-automatic segmentation methods are very time consuming and can show high inter and intra-rater variability [5]. To solve this issue, we present a fully automated method for MS lesions segmentation based on the combination of intensity standardization and voxel-wise comparison of multi-channel Magnetic Resonance Images (MRI) of the patient and control subjects. Our process is applied to the FLI 2016 MSSEG challenge data.

2 Challenge data and evaluation criteria

2.1 Data and pre-processing

To allow challengers to optimize their segmentation algorithms, MS-SEG challenge organizers gave access to 15 MS patient data sets. Each data set contains pre-processed and unprocessed data, available for challengers who wish to perform their own pre-processing on the data, along with the ground truth and the seven manual segmentations used to compute it.

The challenge data sets include T1-w, T1-w Gadolinium, T2-w, PD and FLAIR sequences[1]. Pre-processed data are provided in order to reduce the dependency of the segmentation results on pre-processing performance. The pre-processed data are denoised with the NL-means algorithm [4], rigidly registered [2] towards the FLAIR images, brain extracted using the volBrain platform [8] and bias corrected using the N4 algorithm [12]. As brain extraction was performed, brain masks are provided in the pre-processed data sets. We decided to use the pre-processed data. For this reason, the method we describe below will focus only on the MS lesions segmentation itself.

2.2 Evaluation criteria for the MS-SEG challenge

The quality of the proposed segmentation algorithm may be assessed through two categories of evaluation metrics: lesion detection (are the lesions well detected independently of the contour quality?) and segmentation precision (are the lesion contours close to those of the ground truth?). For the MS-SEG challenge, the organizers will compare the results of the different segmentation workflows to a ground truth for each MS patient and will use several evaluation metrics, out of which two will be used for the ranking of the challengers algorithms[2]:

F1 score: This metric is used to evaluate the quality of an algorithm in terms of lesions detection. It corresponds to the combination of the lesion sensitivity (SensL), i.e. the proportion of detected lesions in the ground truth, and the lesion positive predictive value (PPVL), i.e. the proportion of true positive lesions inside the result of segmentation algorithms.

Dice score: This well known overlap metric is used to evaluate the quality of an algorithm in terms of segmentation precision.

The ground truth is computed with the Logarithmic Opinion Pool based STAPLE (LOP STAPLE) method [1], using seven independent manual segmentations for each patient.

3 MS lesions segmentation workflow

3.1 Intensity Standardization

Our segmentation workflow is based on the voxel-wise comparison of MS patient images against a set of controls. However, intensity profile of conventional MRI has a high inter-subject and inter-scanner variability. To solve this issue, Karpate et al. [7] proposed the estimation of a correction factor which is used to make corresponding anatomical tissues take on the same intensity profile.

[1] The challenge data sets will not be described further in this paper, more details can be found on the challenge website: `https://portal.fli-iam.irisa.fr/msseg-challenge/data`

[2] More details are provided on the challenge website: `https://portal.fli-iam.irisa.fr/msseg-challenge/evaluation`

Image intensities of a healthy brain can be modeled by a 3-class Gaussian Mixture Model (GMM), where each Gaussian represents one of the brain tissues: White Matter (WM), Gray Matter (GM) and Cerebrospinal Fluid (CSF). MS lesions are considered as outliers of this model. Estimating the three classes parameters is rendered difficult because of MS lesion outlier intensities. We therefore estimate them with a modification of the Maximum Likelihood Estimator (MLE) proposed by Notsu et al. [9], more robust to outliers. This estimation is based on the γ-loss function for the Normal distribution, used to maximize the MLE in the form of γ divergence, and is casted to yield an Expectation Maximisation (EM) algorithm [7].

Once the parameters are estimated, we obtain the means and covariances of tissues for the source and target images. These values are used to define a linear correction function which can be solved by linear regression. The results of the linear regression are then exploited to normalize the intensity profiles of the images.

3.2 MS lesions detection

MRI registration and intensity normalization make lesion segmentation possible through a comparison of vector of intensities between the patient and control subjects. Patient images are registered on the set of controls using a linear registration method, based on the use of a block-matching algorithm as presented in [2,10], and a non linear registration method, based on the estimation of a dense non linear transformation between the images as presented in [11]. The methodology used by Karpate et al. [6] to compare the multi-channel vectors of intensities between MS patient and a group of controls is based on the computation of statistical differences through the Mahalanobis distance [3]. These vectors of intensities are built from the available images and can therefore use any combination of them, like FLAIR, or T2-w and FLAIR, or DP, T2-w and FLAIR, making this a parameter of the algorithm.

3.3 Refinement of the segmentation

The intensities of pixels corresponding to brain tissues can vary in function of the brain region where they are located. Indeed, pixels which belong to the white matter brainstem and cerebellum are usually more intense than pixels which belong to the white matter hemispheres. This phenomenon induces the detection of several false positives in the brainstem and the cerebellum. Therefore, intensity standardization and MS lesions detection are computed on one hand in the two hemispheres and on the other hand in the brainstem and the cerebellum.

3.4 Post-processing

This comparison based segmentation algorithm may generate false positives for several reasons (registration errors, presence of noise in the images ...). Therefore, we add a post-processing step to our segmentation workflow in order to reduce the number of false positives. The post-processing is made of four steps:

1. lesions which have a size lower than 3 mm^3 are removed
2. lesions touching the brain mask border are removed, as they are probably false positives due to vessels or skull stripping errors
3. lesions not sufficiently located in WM are removed, as MS lesions are typically located there
4. lesions which do not touch a mask computed from MS patient T2-w and FLAIR sequences are removed. Lesions are considered as hyper intense in these two modalities, so it is possible to build a mask of "probable lesions", i.e. regions where lesions may appear and out of which no lesion may be seen. This mask is built by automatically thresholding the T2-w and FLAIR images and intersecting those masks. Our segmentation method generates several false positives in the brainstem, therefore, the mask used in this post-processing step also excludes this region.
5. lesions delineations are improved using the mask of "probable lesions" computed in the previous step

4 Results

4.1 MRI sequences used for MS lesions detection

Our algorithm can work with only one MR sequence or with several modalities. We have tested our segmentation workflow with T2-w and FLAIR sequences, discarding T1-w and PD images, as they generally show less MS lesion contrast than T2-w and FLAIR. Table 1 presents an evaluation of our segmentation (without post-processing algorithm) on the 15 training data. We worked on two possibilities to combine T2-w and FLAIR images for MS lesions segmentation:

Intersection: Here, we consider that the T2-w and FLAIR MS patient images are registered and intensity normalized on a set of controls. Vector images are created to combine T2-w and FLAIR sequences, on one hand, from MS patient images, and on the other hand, from control subject images. These vector images are then compared with the method described in section 3.2. Intersection of T2-w and FLAIR is interesting as it generates theoretically less false lesions detection as T2-w or FLAIR segmentation alone. However, this way to combine images has a drawback: in some cases, it is possible that a lesion intensity profile is close to the one of a tissue in a modality, which may reduce the Mahalanobis distance and induce the non-detection of a lesion even if its intensity profile is far from tissues intensity profiles in the other modalities.

Union: Here, we consider that our algorithm has already been used to segment MS patient lesions, first with T2-w sequences, and secondly with FLAIR sequences. Then, both segmentations are added. This induces, in theory, that more lesions are detected than in T2-w and in FLAIR segmentations alone. Yet, this type of combination has an inconvenient: false lesions detections from the both T2-w and FLAIR segmentations are kept, thereby adding some noise in the final segmentation.

	Dice scores	SensL	PPVL	F1 scores
T2-w	0.2610	0.3564	0.0305	0.0540
FLAIR	0.4231	0.5519	0.1021	0.1650
T2-w $\bigcup$ FLAIR	0.2978	0.5554	0.0373	0.0685
T2-w $\bigcap$ FLAIR	0.2476	0.5122	0.0205	0.0390

Table 1: Evaluation of our segmentation algorithm on the training data without post-processing.

4.2 Sample results

Table 2 presents an evaluation of the whole segmentation algorithm with a post-processing step on the training data. We choose to perform segmentation twice, using different MRI sequences each time, in order to remove some false positives and to improve lesions delineation. One process uses FLAIR image when the other uses the intersection of T2-w and FLAIR images as theses modalities are the ones which provide the best results. An example of segmentation result is shown in Figure 1.

	Dice scores	SensL	PPVL	F1 scores
T2-w	0.4010	0.3771	0.1811	0.2102
FLAIR	0.5053	0.5309	0.1803	0.2552
T2-w $\bigcup$ FLAIR	0.4244	0.4932	0.1530	0.2135
T2-w $\bigcap$ FLAIR	0.4396	0.5270	0.1507	0.2166
FLAIR and T2-w $\bigcap$ FLAIR	0.5663	0.4560	0.2871	0.3290

Table 2: Evaluation of our segmentation algorithm with a post-processing step.

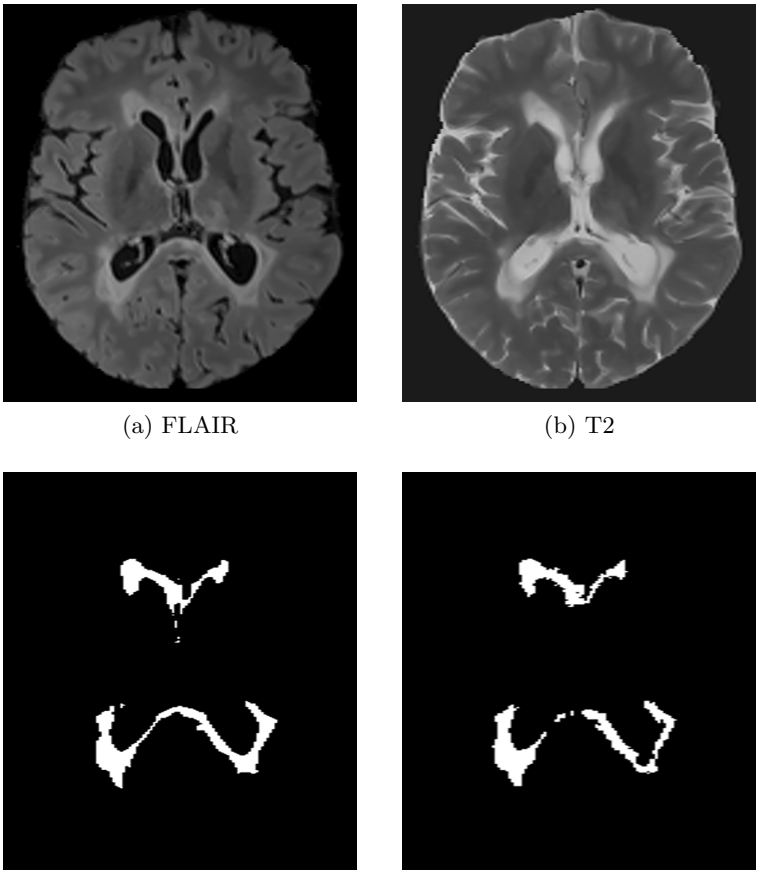
4.3 Implementation and Computation Times

The pipeline presented for the FLI2016 MICCAI MSSEG challenge used the combination of FLAIR and T2 $\bigcap$ FLAIR modalities to perform the segmentation as this process is the one which provides the best results (see Table 2). The algorithm implementation is multi-threaded, based on ITK and available in Anima[3]. The total computation time to process each segmentation of the data set on a computer with an Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz (8 cores) is approximately 8 minutes.

5 Conclusion

We presented a fully automated MS lesion segmentation method based on intensity normalization and voxel-wise comparison. MS lesion segmentation is a

[3] `https://github.com/Inria-Visages/Anima-Public`



(a) FLAIR (b) T2

(c) Ground truth (d) Automatic segmentation

Fig. 1: Automatic segmentation of the data set 01016SACH

complicated task as MS lesion definition is inter-expert dependent. There is a high variability in the detection of lesions even between ground truth and manual segmentations (dice scores and F1 scores for training data vary respectively from 0.26 to 0.88 and 0.13 to 1). A few MR sequences also have a low initial resolution. This may influence the segmentation workflow and lead to worse results. Consequently, the choice of the optimal modalities used to compute a segmentation and the definition of efficient post-processing steps can be a complicated task.

References

1. Akhondi-Asl, A., Hoyte, L., Lockhart, M.E., Warfield, S.K.: A logarithmic opinion pool based STAPLE algorithm for the fusion of segmentations with associated

reliability weights. IEEE transactions on medical imaging 33(10), 1997–2009 (Oct 2014)
2. Commowick, O., Wiest-Daesslé, N., Prima, S.: Block-matching strategies for rigid registration of multimodal medical images. In: 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI). pp. 700–703 (May 2012)
3. Commowick, O., Fillard, P., Clatz, O., Warfield, S.K.: Detection of DTI White Matter Abnormalities in Multiple Sclerosis Patients. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008, pp. 975–982. No. 5241 in Lecture Notes in Computer Science, Springer Berlin Heidelberg (Sep 2008), dOI: 10.1007/978-3-540-85988-8_116
4. Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C.: An Optimized Blockwise Nonlocal Means Denoising Filter for 3-D Magnetic Resonance Images. IEEE Transactions on Medical Imaging 27(4), 425–441 (Apr 2008)
5. Grimaud, J., Lai, M., Thorpe, J., Adeleine, P., Wang, L., Barker, G.J., Plummer, D.L., Tofts, P.S., McDonald, W.I., Miller, D.H.: Quantification of MRI lesion load in multiple sclerosis: A comparison of three computer-assisted techniques. Magnetic Resonance Imaging 14(5), 495–505 (Jan 1996)
6. Karpate, Y., Commowick, O., Barillot, C.: Robust Detection of Multiple Sclerosis Lesions from Intensity-Normalized Multi-Channel MRI (Feb 2015)
7. Karpate, Y., Commowick, O., Barillot, C., Edan, G.: Longitudinal Intensity Normalization in Multiple Sclerosis Patients. In: Linguraru, M.G., Laura, C.O., Shekhar, R., Wesarg, S., Ballester, M.A.G., Drechsler, K., Sato, Y., Erdt, M. (eds.) Clinical Image-Based Procedures. Translational Research in Medical Imaging, pp. 118–125. No. 8680 in Lecture Notes in Computer Science, Springer International Publishing (Sep 2014), dOI: 10.1007/978-3-319-13909-8_15
8. Manjon, J.V., Coupé, P.: volBrain: An online MRI brain volumetry system. In: Organization for Human Brain Mapping'15. Honolulu, United States (Jun 2015)
9. Notsu, A., Komori, O., Eguchi, S.: Spontaneous Clustering via Minimum Gamma-Divergence. Neural Computation 26(2), 421–448 (Feb 2014)
10. Ourselin, S., Roche, A., Prima, S., Ayache, N.: Block Matching: A General Framework to Improve Robustness of Rigid Registration of Medical Images. In: Delp, S.L., DiGoia, A.M., Jaramaz, B. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000, pp. 557–566. No. 1935 in Lecture Notes in Computer Science, Springer Berlin Heidelberg (Oct 2000), dOI: 10.1007/978-3-540-40899-4_57
11. Suarez, R.O., Commowick, O., Prabhu, S.P., Warfield, S.K.: Automated delineation of white matter fiber tracts with a multiple region-of-interest approach. NeuroImage 59(4), 3690–3700 (Feb 2012)
12. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C.: N4ITK: Improved N3 Bias Correction. IEEE Transactions on Medical Imaging 29(6), 1310–1320 (Jun 2010)