



DOCTORAL THESIS

**Automated brain tissue segmentation of
Magnetic Resonance Images in Multiple
Sclerosis**

Sergi Valverde

2016



DOCTORAL THESIS

Automated brain tissue segmentation of Magnetic Resonance Images in Multiple Sclerosis

Sergi Valverde

2016

DOCTORAL PROGRAM in TECHNOLOGY

Supervised by:
Dr. Arnau Oliver
Dr. Xavier Lladó

Work submitted to the University of Girona in partial fulfillment of
the requirements for the degree of Doctor of Philosophy

Publications

The presented thesis is a compendium of the following research articles:

- **Sergi Valverde**, Eloy Roura, Arnau Oliver, Eloy Roura, Sandra González, Deborah Pareto, Joan C. Vilanova, Lluís Ramió-Torrentà, Àlex Rovira and Xavier Lladó. Automated brain tissue segmentation of MR images in the presence of white matter lesions. *Submitted to Medical Image analysis*. 2016. Elsevier. Quality index: [JCR CSAI IF 3.654, Q1 (10/121)].
- **Sergi Valverde**, Arnau Oliver, Eloy Roura, Deborah Pareto, Joan Carles Vilanova, Lluís Ramió-Torrentà, Jaume Sastre-Garriga, Xavier Montalban, Àlex Rovira and Xavier Lladó. Quantifying brain tissue volume in multiple sclerosis with automated lesion segmentation and filling SPM8 toolboxes. *NeuroImage: Clinical*. Vol. 9, pp 640-647. 2016. Elsevier. Quality index: [JCR N IF 2.526, Q2(6/14)].
- **Sergi Valverde**, Arnau Oliver, Yago Díez, Mariano Cabezas, Joan Carles Vilanova, Lluís Ramió-Torrentà, Àlex Rovira, and Xavier Lladó. Evaluating the effects of white matter multiple sclerosis lesions on the volume estimation of six brain tissue segmentation methods. *American Journal of Neuroradiology*. Vol. 36(6), pp. 1109-1115. 2015. American Society of Neuroradiology. Quality index: [JCR RNMMI IF 3.589, Q1(19/125)].
- **Sergi Valverde**, Arnau Oliver, Mariano Cabezas, Eloy Roura and Xavier Lladó. Comparison of ten brain tissue segmentation methods using revisited IBSR annotations. *Journal of Magnetic Resonance Imaging*. Vol. 41, Issue 1, pp. 93-101. January 2015. Wiley. Quality index: [JCR RNMMI IF: 3.210 Q1(23/125)].
- **Sergi Valverde**, Arnau Oliver, and Xavier Lladó. A white matter lesion-filling approach to improve brain tissue volume measurements. *NeuroImage: Clinical*. Vol. 6, pp 86-92. 2014. Elsevier. Quality index: [JCR N IF 2.526, Q2(6/14)].

The rest of publications and conferences related with this PhD thesis are the following:

Journals

- Eloy Roura, Nicolae Sarbu, Arnau Oliver, **Sergi Valverde**, Sandra González, Ricard Cervera, Nuria Bargalló and Xavier Lladó. Automated detection of Lupus white matter lesions in MRI images. *Submitted to Frontiers in Human Neuroscience*. 2016. Frontiers. Quality index: [JCR PS IF 3.626, Q1(13/76)].
- Eloy Roura, Arnau Oliver, Mariano Cabezas, **Sergi Valverde**, Deborah Pareto, Joan Carles Vilanova, Lluís Ramió-Torrentà, Àlex Rovira and Xavier Lladó. A toolbox for multiple sclerosis lesion segmentation. *Neuroradiology*. Vol. 57(10), pp. 1031-1043. 2015. Springer. Quality index: [JCR RNMNI IF 2.485, Q2(41/125)].
- Mariano Cabezas, Arnau Oliver, **Sergi Valverde**, Brigitte Beltrán, Joan Carles Vilanova, Lluís Ramió-Torrentà, Àlex Rovira and Xavier Lladó. BOOST: a supervised approach for multiple sclerosis lesion segmentation. *Journal of Neuroscience Methods*. Vol. 237, pp 108-117. 2014. Elsevier. Quality index: [JCR N IF 2.025, Q3(174/252)].
- Yago Díez, Arnau Oliver, Mariano Cabezas, **Sergi Valverde**, Robert Martí, Joan Carles Vilanova, Lluís Ramió-Torrentà, Àlex Rovira and Xavier Lladó. Intensity based methods for brain MRI longitudinal registration. A study on multiple sclerosis patients. *Neuroinformatics*. Vol 12(3), pp 365-379. 2014. Springer. Quality index: [JCR CSTM IF 2.825, Q1(13/102)]

Conferences

- Eloy Roura, Arnau Oliver, Mariano Cabezas, **Sergi Valverde**, Deborah Pareto, Joan Carles Vilanova, Lluís Ramió-Torrentà, Àlex Rovira and Xavier Lladó. An SPM12 extension for multiple sclerosis lesion segmentation. *SPIE Medical Imaging 2016*. February 2016, San Diego, USA.
- **Sergi Valverde**, Arnau Oliver, Eloy Roura, Deborah Pareto, Joan Carles Vilanova, Lluís Ramió-Torrentà, Jaume Sastre-Garriga, Xavier Montalban, Àlex Rovira and Xavier Lladó. Evaluation of two automated lesion segmentation and filling pipelines for brain tissue segmentation of multiple sclerosis patients.

ECTRIMS 2015. Multiple Sclerosis. October 2015, Barcelona, Spain. Quality index: [JCR CN IF:4.472 Q1(25/191)].

- Eloy Roura, Arnau Oliver, Mariano Cabezas, **Sergi Valverde**, Deborah Pareto, Joan Carles Vilanova, Lluís Ramió-Torrentà, Àlex Rovira and Xavier Lladó. A toolbox for segmenting multiple sclerosis lesions using T1w and FLAIR images. *ECTRIMS 2015. Multiple Sclerosis.* October 2015, Barcelona, Spain. Quality index: [JCR CN IF:4.472 Q1(25/191)].
- **Sergi Valverde**, Arnau Oliver, Deborah Pareto, Joan Carles Vilanova, Àlex Rovira , Lluís Ramió-Torrentà and Xavier Lladó. SLF: a MS white matter lesion filling toolbox for the SPM software. *ECTRIMS 2014. Multiple Sclerosis.* September 2014, Boston, USA. Quality index: [JCR CN IF:4.822 Q1(22/192)].
- Ester Quintana, Brigitte Beltrán, **Sergi Valverde**, René Robles-Cedeno, Héctor Perkal, Xavier Lladó, José Manuel Fernández-Real and Lluís Ramió-Torrentà. Expression of miRNAs in multiple sclerosis cerebrospinal fluid and their relation to MR activity. *ECTRIMS 2014. Multiple Sclerosis.* September 2014, Boston, USA. Quality index: [JCR CN IF:4.822 Q1(22/192)].
- Ester Quintana, Brigitte Beltrán, **Sergi Valverde**, René Robles-Cedeno, Héctor Perkal, Xavier Lladó, José Manuel Fernández-Real and Lluís Ramió-Torrentà. Relación entre la expresión de mirnas en LCR y variables de RM. *Neurología*, vol 29, pp 66-67. 2014. Quality index: [JCR CN IF:1.322 Q3(142/191)].
- **Sergi Valverde**, Arnau Oliver, Mariano Cabezas, Yago Díez, Jordi Freixenet, Xavier Lladó, Joan Carles Vilanova, Àlex Rovira and Lluís Ramió-Torrentà. A quantitative study of the effects of White Matter MS lesions on tissue segmentation methods. *ECTRIMS 2013. Multiple Sclerosis.* October 2013, Copenhagen, Denmark. Quality index: [JCR CN IF:4.472 Q1(25/191)].

Acronyms

- ANN** Artificial Neural Network
BET Brain Extraction Tool
BSE Brain Surface Extractor
CNS Central Nervous System
CSF Cerebrospinal Fluid
CIS Clinically Isolated Syndrome
EDSS Extended Disability Status Scale
FANTASM Fuzzy and Noise Tolerant Adaptive Segmentation Method
FAST FMRIB's Automated Segmentation Tool
FCM Fuzzy-C Means
FLAIR Fluid Attenuated Inversion Recovery
FMRIB Oxford Centre for Functional MRI of the Brain
FSL FMRIB Software Library
GAMIXTURE Image segmentation toolbox based on genetic algorithm and mixture model optimization
GM Gray Matter
IBSR Internet Brain Segmentation Repository
LST Lesion Segmentation Toolbox
MARGA Multispectral Adaptive Region Growing Algorithm
MRI Magnetic Resonance Image
MRBrainS13 Magnetic Resonance Brain Segmentation Challenge 2013
MS Multiple Sclerosis
KNN K-Nearest Neighbor
PD Proton Density
PVC Partial Volume Classifier
SLS SALEM Lesion Segmentation
SPVASEG Image segmentation toolbox based on local Markov random fields
SPM Statistical Parametric Mapping
T1-w T1-weighted
T2-w T2-weighted
WM White Matter

List of Figures

1.1	MRI image modalities	3
1.2	MRI pre-processing steps	4
1.3	Lesion filling example on a T1-w sequence	6
1.4	Organization of the document	11

Abstract

Nowadays, Multiple Sclerosis (MS) is the most frequent non-traumatic neurological disease that causes more disability in young adults. MS is the most common chronic immune-mediated disabling neurological disease of the Central Nervous System, in which the insulating covers of the nerve cells in the spinal chord and brain are damaged. MS is characterized by the presence of lesions in the brain, predominantly in the white matter (WM) tissue of the brain. Due to the sensitivity of Magnetic Resonance Imaging (MRI) to reveal focal WM lesions and disease activity in time and space, MRI has become an essential tool for the diagnosis and evaluation of the MS disease. Furthermore, MRI brain tissue atrophy measures have been shown to correlate with the disability status, showing that tissue loss is an important component of the disease progression.

The existent correlation between brain tissue atrophy measures and MS disability status has increased the necessity to develop robust automated brain tissue segmentation methods capable to perform accurate brain tissue volume measurements. However, automated segmentation of brain tissue is still a challenging problem due to the complexity of the images, differences in tissue intensities, noise, intensity inhomogeneities and the absence of models of the anatomy that fully capture the possible deformations in each structure. Moreover, it has been shown that WM lesions reduce the accuracy of automated tissue segmentation methods, which highlights the necessity to process these lesions before tissue segmentation, a process which is known as lesion filling. However, lesion filling requires to manually annotate lesions before tissue segmentation, which is time-consuming, prone to variability between expert radiologists, or not always available. This fact and the need to analyze quantitatively focal MS lesions in individual and temporal studies has led to the development of a wide number of automated lesion segmentation of MS lesions.

The main goal of this thesis is to develop a novel fully automated brain tissue segmentation method capable of computing accurate tissue volume measurements on images of MS patients containing lesions. In order to fulfill this goal, in this thesis we have focused on each of the concatenated processes that are necessary to develop a fully automated tissue segmentation method. Firstly, we have analyzed and

evaluated the state-of-the-art of tissue segmentation methods on data from healthy subjects, where we have performed a quantitative review of the different proposed tissue segmentation techniques, with the aim to understand their advantages and drawbacks. Our experimental results have shown that methods that incorporate morphological prior information and/or spatial constraints are less prone to changes in acquisition sequences and intensity inhomogeneities, when compared with simpler strategies intensity based methods.

In a second stage, we have studied and evaluated the effect of WM lesions on tissue segmentation of MS patient images. In this regard, we have performed several experiments using multi-center 1.5T MS data from different scanners in order to analyze the effects of lesion signal intensity and lesion size on the differences in tissue volume of several tissue segmentation methods. In all methods, obtained results have indicated that the inclusion of WM lesions on tissue segmentation not only biased the total tissue volume measurements by the addition of miss-classified lesion voxels, but also had a direct effect on the observed differences in normal-appearing tissue. This effect has been less relevant in those methods that incorporate prior information and/or spatial context.

In a third stage, we have focused on lesion filling, reviewing and analyzing the accuracy of different proposed lesion filling techniques in the literature. From this results, we have proposed a new lesion filling technique with the aim to overcome the limitations of previously proposed methods. When compared with these methods, our experimental results have shown that the proposed lesion filling method is effective with different databases and independently of the tissue segmentation method used afterwards.

Afterwards, we have focused on a comprehensive analysis of the effects of automated lesion segmentation and filling in tissue segmentation. We have evaluated the accuracy of two pipelines that incorporated automated lesion segmentation, lesion filling and tissue segmentation on MS data, with the aim to understand the extend of the effect of remaining WM lesions on the differences in tissue segmentation. Our findings have evidenced that up to certain lesion load, pipelines that incorporated automated lesion segmentation and filling are capable to reduce significantly the impact of WM lesions on tissue segmentation, showing a similar performance to the pipelines where expert lesion annotations were used.

All these stages has served as a basis to develop a novel multi-channel method designed to segment brain tissues in MRI images of MS patients. The proposed tissue segmentation method has been designed and implemented using a combination of intensity, anatomical and morphological prior maps to guide the tissue segmentation. WM outliers have been estimated and filled before segmentation using a multi-channel post-processing rule-based algorithm using spatial context, and prior anatomical and morphological atlases. The proposed method has been quan-

titatively and qualitatively evaluated using different databases of images containing WM lesions, yielding competitive and consistent results in both general and MS specific databases. The percentages of error obtained in the different experiments carried out show that the proposed algorithm effectively improves automated brain tissue segmentation in images containing lesions.

This PhD thesis is part of several existing project frameworks carried out by our research group in collaboration with different hospital centers. As part of the goals of these research projects, software implementations of all the proposed methods in this thesis have been released for public use of the research community. The proposed lesion filling method is currently being used by the collaborating hospitals. We believe that the proposed fully automated tissue segmentation method will be also beneficial in clinical settings.

Contents

Publications	i
Acronyms	v
List of figures	vii
Abstract	ix
1 Introduction	1
1.1 Research context	1
1.1.1 Multiple Sclerosis	1
1.1.2 Magnetic Resonance Imaging in MS	2
1.1.3 Image analysis in MS	3
1.2 Research background	6
1.3 Objectives	8
1.3.1 Document structure	10
2 Comparison of 10 brain tissue segmentation methods using revisited IBSR annotations.	13
3 Evaluating the effects of white matter multiple sclerosis lesions on the volume estimation of 6 brain tissue segmentation methods.	15
4 A white matter lesion-filling approach to improve brain tissue volume measurements.	17

5 Quantifying brain tissue volume in multiple sclerosis with automated lesion segmentation and filling.	19
6 Automated tissue segmentation of MR brain images in the presence of white matter lesions.	21
7 Main results and discussion	45
7.1 Effect of WM lesions on tissue segmentation	45
7.2 Effect of lesion filling in tissue segmentation	47
7.3 Effect of automating lesion segmentation and filling on tissue segmentation	48
7.4 Fully automated tissue segmentation of images containing WM lesions	49
8 Conclusions	53
8.1 Future work	55

Chapter 1

Introduction

In this first chapter, we first introduce the reader to the research context and background of the thesis, situating this work within the research line of our group. Afterwards, we describe the proposed objectives and the respective stages to cover. Finally, we summarize the main structure of this thesis, highlighting the conceptual thread between each of the articles that compose its main core.

1.1 Research context

1.1.1 Multiple Sclerosis

The human nervous system can be divided into the central nervous system (CNS) consisting on the brain and the spinal chord, and the peripheral nervous system, which connects the CNS with the sense organs [7]. CNS is mainly constituted by two tissue components: gray matter (GM), which consists of neuronal cell bodies; and white matter tissue (WM), which is mainly composed of myelinated axon tracts [63]. In the case of the brain, it is mostly composed by GM and WM, both evolved by the Cerebro-spinal fluid (CSF), which provides basic mechanical and immunological protection to the brain inside the skull [63].

Multiple sclerosis (MS) is the most common chronic immune-mediated disabling neurological disease of the CNS [64], in which the insulating covers of the nerve cells in the spinal chord and brain are damaged [18]. Nowadays, MS is the most frequent non-traumatic neurological disease that causes more disability in young adults. It follows a similar behavior also seen in other putative autoimmune diseases, and affects twice as many women as it does in men [19]. It has a low incidence in childhood, but it increases rapidly in adulthood reaching a peak between 25 and 35 years, and then slowly declines, becoming rare at 50 and older [8]. So far, the world

estimate for the disease is between 1.3 to 2.5 million cases, being relatively common in Europe, the United States, Canada, New Zealand, and parts of Australia, but rare in Asia, and in the tropics and subtropics [8].

MS is characterized by areas of inflammation, demyelination, axonal loss, and gliosis scattered throughout the CNS, often causing motor, sensorial, vision, coordination, deambulation, and cognitive impairment [17]. Demyelination is the process of progressive damage to the protective covering (myelin sheath) around the axon of the neurons. Demyelinated axons conduct impulses at reduced or spontaneous velocity causing impairment in sensation, movement and cognition [18]. The different clinical courses of the disease are generally grouped into four subtype forms [43]. The *Relapsing/Remitting* (RRMS) form of the disease is characterized by exacerbation times where symptoms are present. These periods are followed by periods of remission, where the patient recovers partial or totally from the disease symptoms. The *Secondary Progressive* (SPMS) form is characterized by a gradual intensification of symptoms between affection relapses. The *Progressive remitting* (PRMS) form is typified by an increase in the relapsing times with significant recovery but with worsening symptoms in new relapsing intervals. Lastly, the *Primary Progressive* (PPMS) form is characterized by a severe decrease of remitting times with special localization in the brain. In general, 50% of RRMS patients after 10 years develop the SPMS form of the disease. After 25 years, the 90% of RRMS patients would develop the SPMS form [43].

1.1.2 Magnetic Resonance Imaging in MS

Magnetic Resonance Imaging (MRI) is a noninvasive medical imaging technique that is used in radiology to generate image representations of different internal anatomical organs and physiological processes of the body. Over the last 40 years, MRI have evolved as a clinical modality [33], and in particular as an essential tool for the diagnosis and evaluation of central nervous system disorders such as MS [24]. On MRI, MS plaques are well-delimited regions with hypointense signal intensity with respect to GM on T1-weighted (T1-w), while hyperintense with respect to GM on T2-weighted (T2-w), Proton Density-weighted (PD-w) and Fluid Attenuated Inversion Recovery (FLAIR) modalities (see Figure 1.1).

In this aspect, new criteria for MS diagnosis and monitoring has been revised in the last years [48], due to the MRI sensitivity to reveal focal white matter (WM) lesions and disease activity in time and space [26]. Additionally, various studies have also analyzed the correlation between MRI brain tissue atrophy measures and MS disability status, showing that tissue loss is an important component of the disease progression [14, 25, 27, 56]. Tissue loss seems to increase through the course of MS with a similar rate between 0.3% and 0.5% per year, and independently of the MS

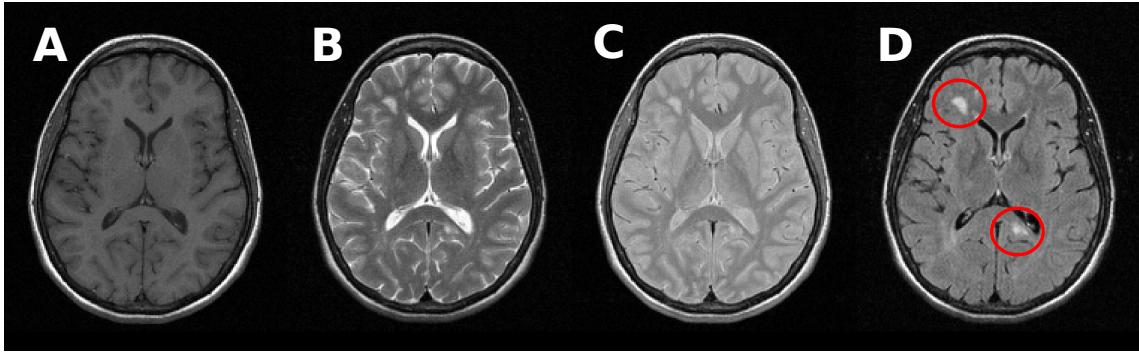


Figure 1.1: MRI image modalities. A) T1-weighted (T1-w) image sequence. B) T2-weighted (T2-w) image sequence. C) Proton Density-weighted (PD-w) image sequence. D) Fluid Attenuated Inversion Recovery (FLAIR) sequence. MS plaques are shown inside red circles on the FLAIR modality. MS plaques are hypointense with respect to GM and WM in T2-w, PD-w and FLAIR sequences, while hypointense with respect to WM on the T1-w modality.

subtype [21, 56]. In general, GM atrophy is more associated with disability changes than WM atrophy [28], and not only in the RRMS and SPMS MS subtypes [27, 56], but also in Clinically Isolated Syndrome (CIS) patients where several studies have shown a significantly greater ventricular cavities and an associated GM loss on MRI scans of CIS patients that will develop MS compared to those who not [13, 25].

1.1.3 Image analysis in MS

Manual analysis of brain images is unfeasible in practice, given the large number of two-dimensional slices of each three-dimensional MRI patient image and the possible intra/inter observer variability between experts [8]. This has led to the development since the early nineties of a wide number of lesion and tissue segmentation methods, with the aim to reduce the time of manual interaction and the inherent variability of manual annotations [16, 32, 39].

Pre-processing of MRI images

Acquired brain MRI volumes incorporate non-brain tissue parts of the head such as eyes, fat, spinal cord or brain skull. Brain tissue extraction from non-brain tissue is commonly referred in the literature as skull-stripping (see Figure 1.2 B and C). Skull-stripping has a direct effect on the performance of automated methods, as differences in skull stripping would lead into unexpected results in the tissue classification if

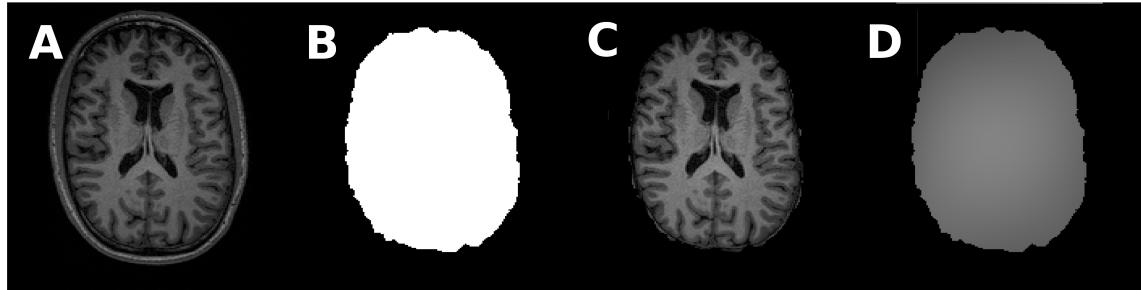


Figure 1.2: MRI pre-processing steps. A) T1-weighted (T1-w) image sequence. B) Computed brain mask using the BET approach [62] and C) skull stripped T1-w sequence . D) Estimated T1-w bias-field using the N3 method proposed by [61].

skull or eyes are included as brain tissue [1, 49]. Among the different proposed methods for skull-stripping [1, 40, 54], the Brain Extraction Tool (BET) [62] and the Brain Surface Extractor (BSE) [59] are the most commonly used methods by the neuroimaging community.

Furthermore, inherent characteristics of the MRI acquisition process such as differences in the magnetic field, bandwidth filtering of the data or eddy currents driven by field gradients usually derive in image artifacts that may also have a negative impact on the performance of methods [60]. In these cases, intensity correction of MRI images is either performed before lesion/tissue segmentation, or as an integrated part of the tissue segmentation pipeline (see Figure 1.2 D). Among the former available strategies proposed [3, 37], the N3 [61] and N4 [68] methods are currently the most widely used tools used for intensity correction.

Automated lesion segmentation

MRI based diagnostic criteria for MS has led to an increasing need to analyze quantitatively focal MS lesions in individual and temporal studies [9, 48]. Different sequences such as T2-w, PD-w and FLAIR are often used in lesion detection and segmentation, as MS lesions appear brighter than GM and WM on them. However, WM lesions often present a similar signal intensity profile to CSF on T2-w. In contrast, FLAIR sequences suppress fluids from the image, restraining the CSF tissue effects on the acquired image, although some severe T2-w hyperintense lesions appear similar to CSF in FLAIR [36].

A wide number of automated lesion segmentation techniques have been proposed during the last years [30, 42]. In these methods, lesion segmentation is based either in supervised or unsupervised strategies. Supervised methods employ a training set of correctly-identified observations that are used as prior information to learn

the lesion characteristics. Newer proposed strategies integrate spatial decision forest [31], statistical methods [65], patch-based models [35] or adaptive dictionary learning strategies [22]. In contrast, unsupervised learning methods do not use any prior information in the segmentation task, which involves grouping data into categories based on some measure of inherent similarity or distance characteristic of the input images. Among these, most recent methods include probabilistic models which separate WM lesions from normal-appearing tissue by considering lesions as an outlier class [36, 38, 67], or techniques that make use of the signal intensity of lesions on FLAIR to apply several thresholding methods with post-processing steps to automatically segment lesions [53, 57].

Automated brain tissue segmentation in MS

The existent correlation between brain tissue atrophy measures and MS disability status [25, 27], has increased the necessity to develop robust automated brain tissue segmentation methods capable to perform accurate brain tissue volume measurements [34]. However, automated segmentation of brain tissue is still a challenging problem due to the complexity of the images, existence of lesions, differences in tissue intensities, noise, intensity inhomogeneities and the absence of models of the anatomy that fully capture the possible deformations in each structure [8, 39].

A wide number of brain tissue segmentation methods have been proposed so far. General purpose intensity based methods usually perform tissue segmentation on T1-w sequences, as this modality clearly separates gray matter from white matter. These include probabilistic strategies based on Bayesian inference [4, 46, 55, 59], Markov Random Fields models [6, 66, 72], or unsupervised clustering methods [11, 47]. In contrast, supervised learning approaches also combine T1-w sequences with other modalities such as T2-w and PD-w using *K-Nearest-Neighbor* classifiers [20, 70], *Support Vector Machines* [2, 69], *Random Forests* [71, 45], or trained *Gaussian mixture models* [51].

However, different studies have shown that tissue abnormalities found in MS image patients such as WM lesions reduce the accuracy of tissue segmentation methods [5, 15]. Effectively, WM lesions on T1-w are hypointense with respect to normal-appearing WM, and therefore, lesion voxels that are classified as GM are distorting the overall GM volume. However, lesion voxels may also have an effect in the observed differences in normal-appearing tissue. WM lesions which are actually classified as WM decrease the mean overall signal intensity of the WM, causing that GM voxels with signal intensities similar to WM lesions may be also miss-classified as WM. In contrast, if WM lesions are classified as GM, normal-appearing WM voxels with signal intensities similar to lesions may be miss-classified as GM.

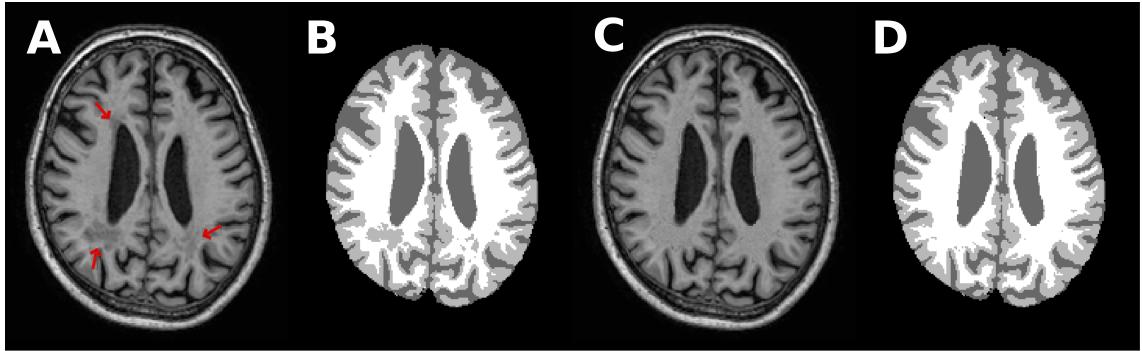


Figure 1.3: Lesion filling example on a T1-w sequence. A) T1-w image sequence containing WM lesions (depicted by red arrows). B) Segmented T1-w sequence containing lesions. GM is depicted in light gray color, WM in white color and CSF in dark gray color. C) T1-w sequence after lesion filling. D) Segmented lesion filled T1-w sequence.

Lesion filling

In MS, hypointense WM lesions have to be pre-processed before tissue segmentation in order to reduce the effects of WM lesions on tissue segmentation. Historically, WM lesions have been masked-out of the T1-w before segmentation, and their volume have been added to WM afterwards [14]. Although this method effectively reduces the error in tissue volume, it has been show in several studies that this approach is not optimal [5, 15].

In this aspect, several strategies have proposed to in-paint lesions on the T1-w with signal intensities of the normal-appearing WM before tissue segmentation [5, 15, 44, 58], a process which is usually known in the literature as lesion filling (see Figure 1.3 for an example). However, most of the available lesion filling methods require manual delineations of lesions, which may be tedious, challenging and time-consuming task depending of the characteristics of the image [42]. When available, lesion filling has demonstrated not only a significant reduction in the associated errors of WM lesions in tissue volume measurements [50], but also in image registration [12, 23, 58] and cortical thickness measurements [44].

1.2 Research background

This thesis is located within the framework of different research projects associated to the Computer Vision and Robotics Institute (VICOROB) of the University of

Girona¹. VICOROB has been working on numerous medical image analysis projects since 1996, mainly in segmentation and registration of mammography images. Lately in 2009, the research group started a fruitful collaboration with several medical MS research teams with the aim to develop new automated techniques capable to segment MS lesions and to perform atrophy measurements that can be transferred to experts for clinical use. In particular, our research in the MS field has been carried out within the following research projects:

1. [2009 – 2012] PI09/91918 “SALEM: Segmentación Automática de Lesiones de Esclerosis Múltiple en imágenes de resonancia magnética” awarded by the Instituto Carlos III.
2. [2009 – 2012] VALTEC09-1-0025 “SALEM: Eines per a la segmentació automàtica de lesions d’Esclerosi Múltiple en ressonància magnètica” awarded in 2009 by the Generalitat de Catalunya within the “Projectes de valorització VALTEC”.
3. [2015–2017] TIN2014-55710-R: NICOLE: “Herramientas de neuroimagen para mejorar el diagnosis y el seguimiento clínico de los pacientes con Esclerosis Múltiple” awarded in 2014 by the spanish call Retos de investigación 2014.
4. [2015 – 2019] BiomarkEM.cat: “New technologies applied to clinical practice for obtaining biomarkers of atrophy and lesions in magnetic resonance images of patients with multiple sclerosis”. Awarded in 2015 by the Fundació la Marató de TV3.

Since then, the research group has published original contributions in different fields such as image pre-processing [54], MS lesion segmentation [9, 10, 42, 53], temporal analysis [29, 41], image registration [23, 52], and tissue segmentation [8]. All the works have been done in collaboration with different medical MS teams from:

- The Hospital Vall d’Hebron: Dr. Rovira, who is the director of the “Unitat de Ressonància Magnètica-Centre Vall d’Hebron” (URMVH) and has participated in numerous research projects funded by public and private institutions in the last few years, as well as Dr. Pareto and technicians Huerga and Corral. This group is part of the MAGNIMS network, a European network of centers that share an interest in the MS study through MRI.
- The Clínica Girona / Hospital Santa Caterina: Dr. Vilanova and Dr. Barceló are the codirectors of the “Unitat de Ressonància Magnètica” at the Clínica Girona and are members of several national and international radiology societies.

¹<http://vicorob.udg.edu>

- The Hospital Josep Trueta: Dr. Ramió-Torrentà, who is the current coordinator of the "Unitat de Neuroimmunologia i Esclerosi Múltiple", as well as Drs. Robles and Beltrán, who work for the radiology unit.

1.3 Objectives

As part of the SALEM, NICOLE and BiomarkEM.cat research project frameworks, the main goal of this thesis is:

to develop a novel fully automated brain tissue segmentation method capable of computing accurate tissue volume measurements on images of MS patients.

Different stages have to be covered first in order to fulfill the main proposed goal. All them can be considered as sub-objectives that allow us to gain a better knowledge of the different parts that compose a fully automated tissue segmentation method for MS images containing lesions. In what follows, we detail these proposed sub-goals:

- **to analyze and evaluate the state-of-the-art of tissue segmentation methods.** This stage aims to quantitative review and evaluate the different proposed tissue segmentation techniques in order to understand their advantages and drawbacks. In order to fulfill this goal, we plan to perform different experiments using public databases that incorporate manual tissue annotations, which will allow to perform a quantitative evaluation of the accuracy of the methods.
- **to study and evaluate the effect of WM lesions on tissue segmentation of MS patient images.** Although it is known that the inclusion of WM lesions on tissue segmentation distort the brain volume measurements, this effect has not been studied and compared across different tissue segmentation methods. In this aspect, the second stage is focused on the analysis of the effects of WM lesions on the tissue distributions of a set of tissue segmentation approaches. Our hypothesis here is that a better knowledge of the correlation between lesion attributes, such as signal intensity and lesion size, and the observed differences in tissue volume of the analyzed algorithms may be beneficial to design a tissue segmentation method for MS. Hence, we aim to perform several experiments using multi-center MS data from different scanners in order to analyze the effects of WM lesions on tissue segmentation.

- **to reduce the effect of WM lesions on tissue segmentation of MS patient images designing and implementing a new lesion filling algorithm.** As said in section 1.1.3, WM lesions have to be pre-processed before tissue segmentation in order to reduce the effects of those lesions on tissue segmentation. In this regard, the third sub-goal is two-fold: firstly, to compare the accuracy of different lesion filling techniques proposed in the literature, analyzing their accuracy on databases with 1.5T and 3T field strengths. Secondly, after analyzing the benefits and drawbacks of each proposed method, we aim to propose a new lesion filling algorithm in order to overcome the possible limitations of existent methods.
- **to analyze and evaluate the effect of automated algorithms that perform WM lesion segmentation and filling on the tissue segmentation.** Although lesion filling techniques have already been successfully applied to reduce the effect of WM lesions on tissue segmentation, usually WM lesions are annotated manually before tissue segmentation. In contrast, the effect of both automated lesion segmentation and filling on tissue segmentation is still unclear. The fourth stage of the thesis aims to understand the effects of the inherent errors in automated lesion segmentation on the posterior lesion filling and tissue segmentation. Thus, we plan to perform several experiments with different pipelines that incorporate automated lesion segmentation, lesion filling and tissue segmentation. Using these experimental data, we aim to evaluate the accuracy of these pipelines on MS data, analyzing and evaluating the extend of the effect of remaining WM lesions on the differences in tissue segmentation, which may be beneficial to update the knowledge gained in previous stages.
- **to propose a new fully automated tissue segmentation method for MS patient images.** Finally, we aim to benefit from the stages to propose a novel **fully automated tissue segmentation method** able to deal with images of MS patients with different level of brain atrophy and lesion load. In this last stage, we aim to validate the accuracy of the proposed method by comparing it with the state-of-the-art in tissue segmentation in MS.

This objective refers to the brain tissue segmentation of MS patient images into GM, WM, CSF in transversal studies. We do not concentrate in differences in tissue volume at different stages, but in the effect of WM lesions in the final tissue segmentation. All these stages will be carried out using not only public databases but also different 1.5T and 3T databases of MS patients from the collaborating hospital centers. **Furthermore, as part of the goals of the research frameworks from which this thesis is located, implementations of all the proposed methods will be publicly available for the research community.**

1.3.1 Document structure

A graphical description of the structure of the thesis linking all the chapters presented is shown in Figure 1.4. Connections between the chapters depict the conceptual link between them. The rest of the document is organized as follows:

- **Chapter 2. Comparison of 10 brain tissue segmentation methods using revisited IBSR annotations.** We present here comprehensive comparison of the accuracy of 10 brain tissue segmentation methods on two public MRI databases. This chapter is based on the paper published in the *Journal of Magnetic Resonance Imaging* in 2015.
- **Chapter 3. Evaluating the effects of white matter multiple sclerosis lesions on the volume estimation of 6 brain tissue segmentation methods.** After reviewing different tissue segmentation techniques using public data, we perform a detailed analysis of the effects of WM lesions on the brain tissue volume measurements of six of these tissue segmentation methods using MS data from different hospital centers collaborating in the research projects. This chapter is based on the paper published in the *American Journal of Neuroradiology* in 2015.
- **Chapter 4. A white matter lesion-filling approach to improve brain tissue volume measurements.** In this chapter, we propose a new technique to fill WM lesions on 1.5T and 3T** TODO Chapter 6: definiciÃ§ del journal Ãls incorrecta data, validating its accuracy with respect to other methods in the literature. This chapter is based on the paper published in the *NeuroImage: Clinical* journal in 2014.
- **Chapter 5. Quantifying brain tissue volume in multiple sclerosis with automated lesion segmentation and filling.** In this chapter we present a detailed evaluation of the performance of different automated pipelines that incorporate lesion segmentation, lesion filling and tissue segmentation on MS data. This analysis is novel in the sense that this is the first work that evaluates two automated pipelines on MS data. This chapter is based on the paper published in the *NeuroImage: Clinical* journal in 2015.
- **Chapter 6. Automated brain tissue segmentation of MR images in the presence of white matter lesions.** We propose here a novel fully automated tissue segmentation pipeline designed to deal with MS patient images containing lesions. We validate the accuracy of the proposed method comparing the performance with other state-of-the-art techniques. Data from the MRBrainS13 challenge as well as data from our hospital collaborators is used

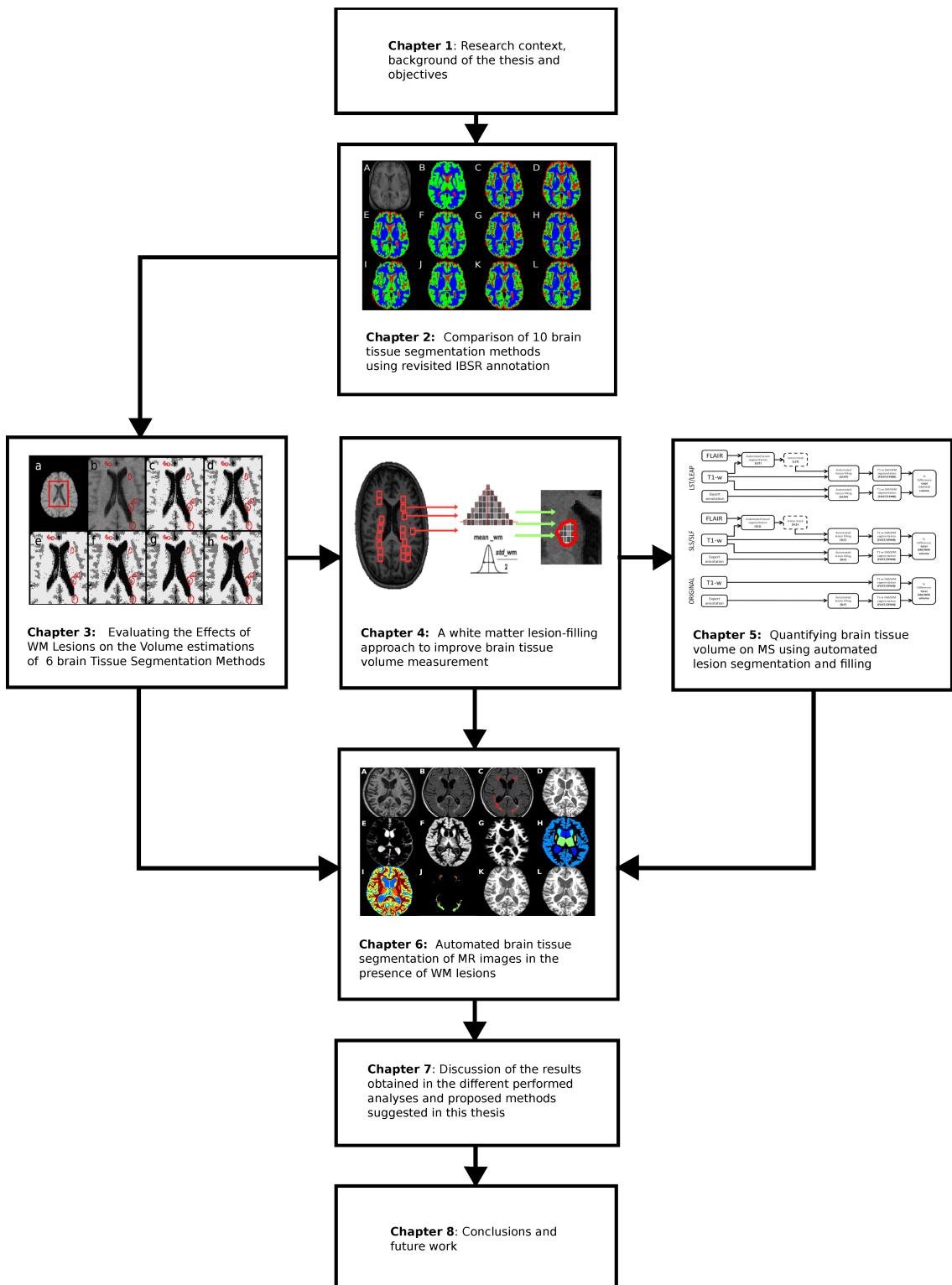


Figure 1.4: Organization of the document. Preliminary chapter 1 describes the research context and the main objectives of this thesis. Chapters 2 to 6 introduce the main contributions of this work based on the different works submitted or published in research journals. Chapter 7 presents a general discussion of the results obtained from Chapters 2 to 6. Finally, the main conclusions and the proposed future work are presented in Chapter 8. Connections between chapters depict a conceptual link between them.

to perform the evaluation. This chapter is based on the paper submitted to the *Medical Image Analysis* journal in 2016.

- **Chapter 7. Results and discussion.** This chapter provides a comprehensive discussion of the results obtained in this thesis.
- **Chapter 8. Conclusions and future work.** Finally, the main conclusions based on the contributions of this thesis are defined. Based on these conclusions, we also point out different future works to improve and extend the work carried out in this thesis.

Chapter 2

Comparison of 10 brain tissue segmentation methods using revisited IBSR annotations.

In this chapter, we perform a quantitative evaluation of the accuracy of 10 automated brain tissue segmentation methods. Methods are compared based using the Internet Brain Segmentation Repository (IBSR) databases IBSR20 and IBSR18¹. The performance of the methods is evaluated by ranking their accuracy based on the significant differences with respect the rest of methods. This proposed evaluation has been published in the following paper:

Paper published in **Journal of Magnetic Resonance in Medicine (JMRI)**

Volume: 41, Issue: 1, Pages: 93-101, Published: January 2015

DOI: 10.1002/jmri.24517

Quality Index: 3.21 (Quartile 1)

¹<https://www.nitrc.org/projects/ibsr/>

Chapter 3

Evaluating the effects of white matter multiple sclerosis lesions on the volume estimation of 6 brain tissue segmentation methods.

In this chapter, we present an study of the impact of MS white matter lesions on the brain tissue measurements of six well-known segmentation techniques. These include straightforward techniques such as Artificial Neural Network (ANN) and fuzzy C-means (FCM) as well as more advanced techniques such as the Fuzzy And Noise Tolerant Adaptive Segmentation Method (FANTASM), FMRIB's Automated Segmentation Tool (FAST), and Statistical Parametric Mapping (SPM) with versions SPM5 and SPM8. This proposed evaluation has been published in the following paper:

Paper published in the **American Journal of Neuroradiology (AJNR)**

Volume: 36, Pages: 1109-1115, Published: February 2015

DOI: 10.3174/ajnr.A4262

Quality Index: 3.59 (Quartile 1)

Chapter 4

A white matter lesion-filling approach to improve brain tissue volume measurements.

In this chapter, we propose a new technique to fill WM lesions before tissue segmentation. The proposed approach is evaluated in both 1.5T and 3T data. We validate our method comparing its accuracy with other proposed automated lesion filling methods on the same data. Furthermore, the proposed technique has been released for public use both as a standalone program or as SPM8/SPM12 library. This work has been published in the following paper:

Paper published in the **NeuroImage: Clinical** journal (NICL)
Volume: 6, Pages: 86-92, Published: August 2014
DOI: [10.1016/j.nicl.2014.08.016](https://doi.org/10.1016/j.nicl.2014.08.016)
Quality Index: 2.53 (Quartile 2)

Chapter 5

Quantifying brain tissue volume in multiple sclerosis with automated lesion segmentation and filling.

In this chapter we present a detailed evaluation of the performance of different pipelines that incorporate fully automated processes such as lesion segmentation, lesion filling and tissue segmentation on MS data. For each automated pipeline, we analyze the percentage of error in tissue segmentation between a set of 70 MS images where WM lesions have been refilled before segmentation and the same images processed different levels of automation from manually masking lesion to fully automated lesion segmentation and filling. This analysis has been published in the following paper:

Paper published in the **NeuroImage: Clinical** journal (NICL)

Volume: 9, Pages: 640-647, Published: October 2015

DOI: doi:10.1016/j.nicl.2015.10.012

Quality Index: 2.53 (Quartile 2)

Chapter 6

Automated tissue segmentation of MR brain images in the presence of white matter lesions.

In this chapter, we propose a novel automated pipeline for tissue segmentation of MS patient images containing lesions. The accuracy of the method is evaluated using both the challenge MRBrainS13 database¹ and a 3T MS database of MS patient images. We validate the accuracy of the proposed method with other state-of-the-art techniques. A public version of the method has been released for public use. The proposed pipeline has been described in detail in the next paper and submitted to the Medical Imaging Journal:

Paper submitted to the **Medical Image Analysis** (MIA) journal
Currently in revision.
Quality Index: 3.68 (Quartile 1)

¹<http://mrbrains13.isi.uu.nl/>

Automated tissue segmentation of MR brain images in the presence of white matter lesions

Sergi Valverde^{a,*}, Arnau Oliver^a, Eloy Roura^a, Sandra González^a, Deborah Pareto^b,
Joan C. Vilanova^c,
Lluís Ramió-Torrentà^d, Alex Rovira^b, Xavier Lladó^a

^aResearch institute of Computer Vision and Robotics, University of Girona, Spain

^bMagnetic Resonance Unit, Dept of Radiology, Vall d'Hebron University Hospital, Spain

^cGirona Magnetic Resonance Center, Spain

^dMultiple Sclerosis and Neuroimmunology Unit, Dr. Josep Trueta University Hospital, Spain

Abstract

During the last years, the increasingly interest of brain tissue volume measurements on clinical settings has lead to the development of a wide number of automated tissue segmentation methods. However, white matter lesions are known to reduce the accuracy of automated tissue segmentation methods, which requires to manually annotate lesions and refill them before segmentation, which is tedious and time-consuming. Here, we propose a new fully automated T1-w/FLAIR tissue segmentation approach designed to deal with images in the presence of WM lesions, which integrates a robust partial volume tissue segmentation with WM outlier rejection and filling, combining intensity, probabilistic and morphological prior maps. We evaluate the accuracy of the method on the MRBrainS13 tissue segmentation challenge database, and also on a set of Multiple Sclerosis (MS) patient images. On both databases, we validate the performance of our method with other state-of-the-art techniques. On the MRBrainS13 data, the presented approach was the best unsupervised ranked method of the challenge (7th position) and clearly outperformed other unsupervised pipelines such as *FAST* and *SPM12*. On MS data, the differences in tissue segmentation between the images segmented with our method and the same images where manual expert annotations were used to refill lesions on T1-w images before segmentation were lower or similar to the best state-of-the-art pipeline incorporating automated lesion segmentation and filling. Our results show that the proposed pipeline quantitatively improved the accuracy of tissue segmentation while it achieved very competitive results on MS images. A public version of the approach is available to download for the neuro-imaging community.

Keywords: Brain, MRI, multiple sclerosis, automatic tissue segmentation, white matter lesions

1. Introduction

Brain tissue volume based on Magnetic Resonance Imaging (MRI) is increasingly being used in clinical settings to assess brain volume in different neurological diseases such as the Multiple Sclerosis (MS) (Giorgio and De Stefano, 2013). In MS, several studies have analyzed the histopathological changes of patients with respect to the progression of the disease, showing that the percentage of change in brain volume tends to correlate with worsening conditions (Pérez-Miralles et al., 2013; Sormani et al., 2014). However, manual segmentation of brain tissue is both challenging and time-consuming because of the large number of MRI slices for each patient which composes the three-dimensional information,

*Corresponding author. S. Valverde, Ed. P-IV, Campus Montilivi, University of Girona, 17071 Girona (Spain). e-mail: svalverde@eia.udg.edu. Phone: +34 972 418878; Fax: +34 972 418976.

and the inherent intra/inter-observer variability of manually segmented scans (Cabezas et al., 2011). The development of automated MS tissue segmentation methods that can segment large amounts of MRI data, do not suffer from intra/inter-observer variability and specific changes of the brain such as MS associated lesions and brain atrophy are still an active research field (Klauschen et al., 2009; de Bresser et al., 2011; Valverde et al., 2015a; Vincken et al., 2015).

Different brain tissue segmentation methods have been used in MS so far. General purpose intensity based methods combining intensity and a priori statistical anatomic information such as FAST (Zhang et al., 2001) or SPM (Ashburner and Friston, 2005) are nowadays widely used. However, tissue abnormalities found in MS image patients such as White Matter (WM) lesions reduce the accuracy of these techniques (Chard et al., 2010; Battaglini et al., 2012) usually causing an overestimation of Gray Matter (GM) tissue not only by the effect of hypointense WM lesion voxels classified as GM, but also by the effect of these lesion voxels in normal-appearing tissue (Valverde et al., 2015b). In these cases, in-painting lesions on the T1-weighted image (T1-w) with signal intensities of the normal-appearing WM before tissue segmentation may be used to reduce the effects of WM lesions on tissue segmentation (Chard et al., 2010; Battaglini et al., 2012; Valverde et al., 2014). However, MS lesions have to be delineated manually first, which may be tedious, challenging and time-consuming task depending of the characteristics of the image (Lladó et al., 2012).

Regarding this issue, a wide number of automated lesion segmentation techniques have been proposed during the last years (Lladó et al., 2012; García-Lorenzo et al., 2013). Most of these methods integrate other imaging modalities such as T2-weighted, Proton Density (PD) and Fluid Attenuated Inverse Recovery (FLAIR), as these modalities present a high contrast between tissue and lesions (Lladó et al., 2012). More recent techniques include supervised learning classifiers based on spatial decision forest (Geremia et al., 2011), statistical methods (Sweeney et al., 2013), patch-based models (Guizard et al., 2015) or adaptive dictionary learning methods (Deshpande et al., 2015). Further, different unsupervised learning techniques make use of probabilistic models to separate WM lesions from normal-appearing tissue by considering lesions as an outlier class (Harmouche et al., 2015; Tomas-Fernandez and Warfield, 2015; Jain et al., 2015). Also, other unsupervised techniques make use of the signal intensity of lesions on FLAIR several thresholding methods with post-processing steps to automatically segment lesions (Schmidt et al., 2012; Roura et al., 2015). In contrast, there are fewer studies that have been focused on the tissue segmentation of MS images containing lesions. Those include non-supervised techiques combining intensity, anatomical and morphological maps (Nakamura and Fisher, 2009; Shiee et al., 2010), or supervised methods such as statistical classifiers (Datta and Narayana, 2013), atlas based nearest-neighbor methods (De Boer et al., 2009) and sparse dictionary learning approaches (Roy et al., 2015).

The increasing amount of published studies regarding automated WM lesion segmentation may be given by the particular need of a quantitative analysis of focal MS lesions in individual and temporal studies (Lladó et al., 2012). Recent studies in MS (Chard et al., 2010; Gelineau-Morel et al., 2012; Ceccarelli et al., 2012; Pérez-Miralles et al., 2013; Popescu et al., 2014; Magon et al., 2014; Valverde et al., 2015b) indicate a certain tendency to the use of widely validated segmentation tools such as Siena (Smith et al., 2002), FAST (Zhang et al., 2001) or SPM (Ashburner and Friston, 2005) in combination with automated lesion segmentation and/or lesion-filling approaches, although their application in clinical practice is still not generalized (Giorgio and De Stefano, 2013).

In this paper, we present the Multiple Sclerosis SEGmentation pipeline (*MSSEG*), a novel multi-channel method designed to segment GM, WM and cerebro-spinal fluid (CSF) tissues

in images of MS patients. This method is motivated by our previous analysis of the effects of tissue segmentation on MS images (Valverde et al., 2015b), the role of lesion-filling (Valverde et al., 2014), and its combination with automated lesion segmentation on tissue segmentation (Valverde et al., 2015b). Similar to the works of Nakamura and Fisher (2009) and Shiee et al. (2010), our approach utilizes a combination of intensity, anatomical and morphological prior maps to guide the tissue segmentation. However, tissue segmentation is here based on a robust partial volume segmentation where WM outliers are estimated and refilled before segmentation using a multi-channel post-processing algorithm. This post-processing algorithm is partially inspired on MS lesion segmentation algorithm proposed by Roura et al. (2015), but here we integrate multi-channel support, partial volume segmentation, spatial context, and prior anatomical and morphological atlases. In order to perform quantitative and qualitative evaluations of our approach, we analyze its accuracy with the MRBrainS13 challenge database which includes manual tissue annotations, and also with a set of MS patient images with different lesion burden. We quantitatively compare the performance of our approach with different state-of-the-art techniques that also competed on the MRBrainS13 challenge and/or have been used in recent MS studies. Furthermore, we also analyze the differences in the performance of our approach when using only T1-w or when using the multi-channel approach that includes T1-w and FLAIR. The *MSSEG* method is currently available for download at our research group webpage (<http://atc.udg.edu/nic/msseg/index.html>).

2. Methods

The proposed brain tissue segmentation method is composed of five different processes: registration of statistical atlas into T1-w space (Sec. 2.2), tissue estimation (Sec. 2.3), detection and re-assignation of lesion candidate to T1-w (Sec. 2.4), tissue re-estimation, and partial volume re-assignation of tissue maps into CSF, GM and WM (Sec. 2.5). The overall schema of the pipeline is depicted in figure 1. We describe each step in detail in the following subsections.

2.1. Notation

To describe our proposed approach, we employ the following notation. T and F denote the input images T1-w and FLAIR, respectively. P^c denotes a probabilistic tissue atlas of particular class $c = \{csf, csfgm, gm, gmwm, wm\}$. S^{st} denotes a morphological brain atlas of a particular parcellated structure st . For each of the above images, T_j , F_j , P_j^c and S_j^{st} denote an observation at a voxel $j \in \Omega$, being Ω the image domain.

2.2. Tissue prior registration

The MNI-ICBM 152 2009a Nonlinear T1-w average structural template image¹ was first affine registered to the native T1-w image space based on a block matching approach (Ourselin et al., 2002), and then non-rigid registered with a fast free-form deformation method (Modat et al., 2010), both using the Nifty Reg package². Obtained transformation parameters were then used to resample the available MNI CSF, GM and WM tissue priors to the T1-w space. The resampled probabilistic tissue maps P^{csf} , P^{gm} and P^{wm} were extended to build intermediate partial volumes P^{csfgm} as $(P^{csf} \geq 0.5 \cap P^{gm} \geq 0.5)$ and P^{gmwm} as $(P^{gm} \geq 0.5 \cap P^{wm} \geq 0.5)$ and taking the mean value of the two input atlases.

¹<http://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009>

²<http://cmictig.cs.ucl.ac.uk/wiki/index.php/NiftyReg>

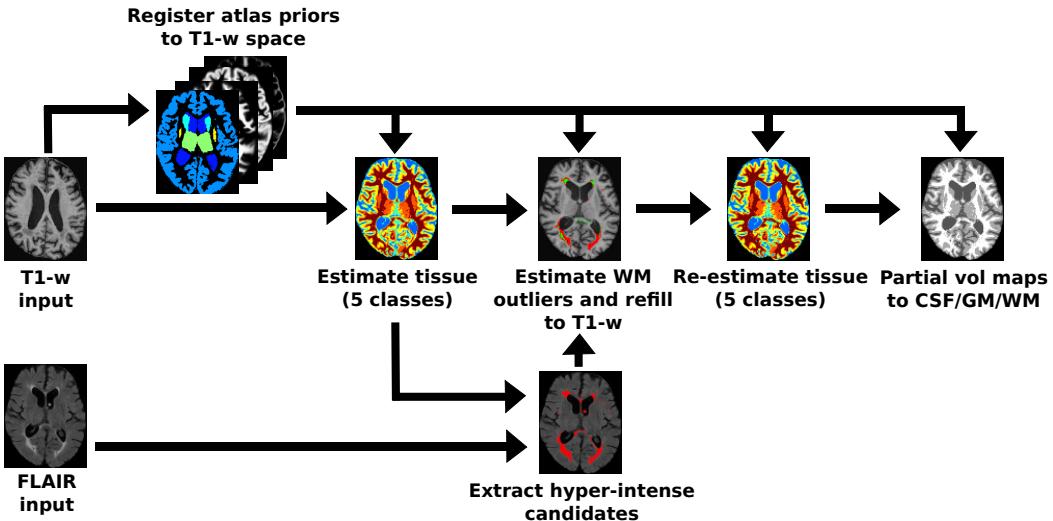


Figure 1: The proposed method *MSSEG* consists of five different steps: 1) Three statistical *a-priori* tissue atlases (CSF, GM and WM) and a brain structure atlas are first registered into the patient space (Sec. 2.2) and then used to 2) guide the tissue segmentation of the input T1-w image (Sec. 2.3). 3) Then, the same output segmentation is employed to detect and reassign candidate regions into WM based on the registered *a-priori* and hyper-intense FLAIR maps if available (Sec. 2.4). The voxel intensities of candidate regions on T1-w are then refilled with normal-appearance WM intensities and 4) tissue is re-estimated (Sec. 2.3). 5) Finally, intermediate volume maps are reassigned into CSF, GM and WM using both neighbor and spatial prior information (Sec. 2.5).

Besides, a morphological brain structure atlas was first parcellated on the original MNI atlas using the hierarchical algorithm proposed by Pohl et al. (2007) on EMSegmenter³, and then manually refined for the selected structures. The resulting atlas (S^{st}) consisted of useful structures for tissue segmentation such as cortical GM (S^{CORTEX}), ventricles (S^{VENT}), basal ganglia (S^{BASAL}) and brainstem ($S^{BRAINSTEM}$). The same transformation parameters were also used to resample the morphological atlas to the T1-w space.

2.3. Tissue estimation

Brain tissue was estimated following a robust fuzzy-clustering approach similar to the one proposed in Pham (2001), as this method provides a straightforward implementation, fairly robust behavior including spatial context information, applicability to multichannel data, and the ability to model uncertainty within the data (Pham, 2001). In our approach, we designed the method to segment 5 classes in order to preserve better the differences in signal intensity between local regions of the brain and lesion candidates. We also extended the spatial penalizing weights by incorporating the probabilistic tissue priors in the segmentation process similarly to Shiee et al. (2010). Hence we modified the objective function proposed by Pham (2001) in order to incorporate also prior-atlas information as follows:

³<https://www.slicer.org/slicerWiki/index.php/EMSegmenter-Overview>

$$\begin{aligned}
J_{MSSEG} = & \sum_{j \in \Omega} \sum_{k=1}^C u_{jk}^q \|T_j - v_k\|^2 + \\
& + \frac{\beta}{2} \sum_{j \in \Omega} \sum_{k=1}^C u_{jk}^q \sum_{l \in N_j^w} \sum_{m \in M_k} u_{lm}^q + \\
& + \frac{\gamma}{2} \sum_{j \in \Omega} \sum_{k=1}^C u_{jk}^q \sum_{l \in N_j^w} \sum_{m \in M_k} P_l^k
\end{aligned} \tag{1}$$

where $\{k \in C \mid C = \{csf, csfgm, gm, gmwm, wm\}\}$, u_{jk} denotes the membership probability of each voxel j for a particular class, v_k are the cluster signal intensity centers of each class, N_j^w is the set of two-dimensional (2D) $(2w+1)^2$ or three-dimensional (3D) $(2w+1)^3$ neighbors centered on the voxel j , and $M_k = \{1, \dots, C\} \setminus \{k\}$. This approach depends on four parameters to adjust the membership functions: the weighting parameter q that controls the degree of fuzziness, the spatial constraint parameter β that controls the amount of neighbor information added, the prior belief parameter γ used to control the amount of prior atlas information about each tissue, and finally the window radius of neighbors w .

Similar to the work of Pham (2001), an iterative algorithm to minimize (1) was derived by evaluating the centroids and the functions that satisfy a zero gradient condition as follows:

Algorithm 1 Tissue estimation

1: Obtain the initial estimates of the centroids for each class $k = \{1, \dots, C\}$:

$$v_k = \frac{1}{n} \sum_{j \in \Omega} (T_j^k \mid P_j^k \geq 0.5) \quad n = |(T_j^k \mid P_j^k \geq 0.5)|$$

2: Compute the membership functions u_{jk}

3: Compute the new centroids:

$$v_k = \frac{\sum_{j \in \Omega} u_{jk}^q T_j}{\sum_{j \in \Omega} u_{jk}^q} \quad k = \{1, \dots, C\}$$

4: Repeat steps 2 and 3 until convergence

Initial centroids v_k were estimated for each class C by taking the mean signal intensity of the voxels on the T1-w image with prior-tissue probability $P_j^k \geq 0.5$. The membership function u_{jk} was also adapted to incorporate prior-atlas information and computed as follows:

$$u_{jk} = \frac{(\|T_j - v_k\|^2 + \beta \sum_{l \in N_j} \sum_{m \in M_k} u_{lm}^q + \gamma \sum_{l \in N_j} \sum_{m \in M_k} P_l^m)^{-1/(q-1)}}{\sum_{i=1}^C (\|T_j - v_i\|^2 + \beta \sum_{l \in N_j} \sum_{m \in M_i} u_{lm}^q + \gamma \sum_{l \in N_j} \sum_{m \in M_k} P_l^m)^{-1/(q-1)}} \tag{2}$$

The five classes tissue segmentation mask SEG_j was computed by assigning to each voxel the class with maximum membership as follows:

$$SEG_j = \arg \max_k u_{jk} \quad \forall j \in \Omega \tag{3}$$

The parameters q , γ and w can be tuned manually to increase the performance of the method, but were set to default values $q = 2$, $\gamma = 0.1$ and $w = 1$ with 2D that worked well in the majority of cases. In contrast, the β parameter depends on the brightness of the image, the deviation of the signal intensities of voxel class members with respect to their centroid value, and image noise (Pham, 2001). Hence, choosing a proper value for the β parameter was important to obtain optimal or near-optimal performance. In our implementation, we automated the choose of the β parameter by fitting a function of the optimal empirical selection of the parameter with respect to different levels of noise. To do so, we estimated iteratively the sub-optimal β parameter of 10 images of the Brainweb

dataset ⁴ that included different noise level (1-9%) and ground-truth annotations. For each image, we also computed the noise level using the Fast Noise Variance method proposed by Immerkær (1996). Then, the correspondent β parameters and noise levels were used to fit a polynomial function to interpolate the β parameter. For all the evaluated images in the paper, we have automatically approximated the β as a function $B(x)$ of their noise level x as $B(x) = 0.0011x^4 - 0.0015x^3 + 0.0074x^2 - 0.001x + 0.05$.

2.4. Reassign WM outliers to T1-w

In a three class tissue segmentation approach, lesion regions are usually classified as either GM or WM, given the hypointense signal intensity profile of WM lesions. This impedes in some cases to differentiate them from surrounding GM and WM. In contrast, by creating the intermediate classes *CSFGM* and *GMWM*, new local clusters of voxels with similar signal intensities are delimited, increasing the chances that WM lesion regions may be differentiated from normal-appearing GM and WM. Following this assumption, we estimated WM lesion regions by analyzing all local regions not initially segmented as WM based on their prior probability and the spatial connection to WM.

First, different binary segmentation masks M^c were computed for each of the classes $c = \{gmwm, gm, csfgm\}$. For each mask, all 2D regions of connected components were computed using a flood-fill algorithm with 4-connected neighborhood. We define the set of all 4-connected p regions given an input binary image as follows:

$$R_p^T \leftarrow \oplus(M_j^c, n), \quad p = \{1, \dots, |R_p^T|\}$$

where the operator \oplus refers to the connected components function and n is the number of connected neighbors.

Secondly, a map of hyperintense region candidates was computed on the FLAIR image following the same strategy shown in Roura et al. (2015). The binary mask M^{GM} was first used to compute the intensity distribution on the FLAIR image, where GM is typically hyperintense with respect to CSF and WM, and WM lesions are considered hyperintense outliers to GM. The mean and standard deviation of the GM distribution was computed using the full-width at half maximum (FWHM) of the main peak of a generated histogram. Then, an initial map of hyperintense regions voxels H^{FLAIR} was determined by thresholding the FLAIR image F as follows:

$$H_j^{FLAIR} = \begin{cases} 1 & \text{if } F_j > \mu + \alpha\sigma \\ 0 & \text{otherwise} \end{cases} \quad \forall j \in \Omega \quad (4)$$

where μ and σ were the mean and standard deviation, respectively, of the GM distribution as computed using the FWHM, and α was a weighting parameter that scaled the minimum signal intensity of outliers. The binary mask H^{FLAIR} was then used to group the candidate voxels into connected regions using the same method proposed before:

$$R_t^F \leftarrow \oplus(H^{FLAIR}, n), \quad t = \{1, \dots, |R_t^F|\}$$

where n was set to 3D connected elements ($n = 6$) in order to reduce the amount of 2D false positive regions such as hyperintense sub-arachnoid tissue.

Given the computed binary masks for each tissue M^{gmwm} , M^{gm} and M^{csfgm} , the map of hyperintense voxels on FLAIR H^{FLAIR} , and its connected components R_t^F , we used an

⁴<http://brainweb.bic.mni.mcgill.ca/brainweb/>

iterative algorithm to estimate the regions with high probability to pertain to WM. Lesion filling of selected regions was integrated in the same algorithm. Figure 2 shows in detail each of the steps of the algorithm. Regions not overlapping cortical GM on the morphological prior S_s^{CORTEX} were only processed if a matched region was also hyperintense in FLAIR, in order to reduce the amount of false positive regions such as isolated cortical GM segmented regions. Regions not touching cortical GM were filtered based on their prior probability to pertain to WM and their distance to surrounding WM. If half of the voxels of a region had a prior probability to pertain to WM and the region was connected to actual WM, the region was also added to the WM class, and those voxels were refilled as normal-appearing WM into the original T1-w image T using the same implementation proposed in Valverde et al. (2014). Note that classes were visited from $gmwm$ to $csfgm$ in order to add new belief of the actual WM and use it to filter next regions.

If the FLAIR modality is not available, H^{FLAIR} is automatically set to zero, disabling the evaluation of R^H regions and henceforth forcing the method to evaluate the next T1 region R_p^T . If FLAIR is used, all regions that were discarded on the first part of the algorithm or overlapped the cortex, were filtered according to their spatial attributes on the FLAIR image. Each discarded region R_q^T on the segmented mask SEG was matched with a particular region in FLAIR R_t^H based on their overlap ($R_t^H \mid t = \arg \max_t (|R_t^H \cap R_q^T|)$). Then, matched regions where half of their surrounding neighbors were actually classified as $gmwm$ or wm were also added to the WM class and T1-w filled. In all cases, we referred to the neighboring voxels of a region N_S as the neighbors with one voxel of distance from the region boundaries.

2.5. Partial volume maps

Once WM outliers were reassigned to T1-w, the resulting refilled image was used to estimate the brain tissue following the same method described in Sec. 2.3. Afterwards, partial volume maps ($csfgm$) and ($gmwm$) were reassigned to each of the three main classes CSF, GM and WM following a region-wise approach.

Local 2D regions with similar intensity that were classified as $csfgm$ and $gmwm$ were estimated using the same connected component algorithm described before. The structural brain atlas S was then used to reassign regions that at least half of their voxels overlapped with certain structures as follows:

$$SEG_{R_p} = \begin{cases} CSF & \text{if } \left(\frac{1}{|R_p|} \sum_{s \in R_p} S_s^{VENT} \right) > 0.5 \\ GM & \text{if } \left(\frac{1}{2|R_p|} \sum_{s \in R_p} S_s^{CORTEX} + S_s^{BASAL} \right) > 0.5 \\ WM & \text{if } \left(\frac{1}{|R_p|} \sum_{s \in R_p} S_s^{BRAINSTEM} \right) > 0.5 \end{cases} \quad (5)$$

for all the regions $p = \{1, \dots, |R_p|\}$. The rest of voxels not reassigned previously were reclassified by adding them to the surrounding pure class with the most similar intensity as follows:

$$SEG_j = \arg \min_c \left| T_j - \frac{1}{|N_j|} \sum_{s=1}^{|N_j|} (T_s \mid SEG_s = c) \right| \quad (6)$$

for pure classes $c = \{csf, gm, wm\}$ and partial volume voxels $j = \{\forall j \in \Omega \mid SEG_j = csfgm \cup gmwm\}$. The radius for neighbor voxels was set to 6 in two dimensions. Figure 3 depicts the partial volume re-assignment process for one particular T1-w image.

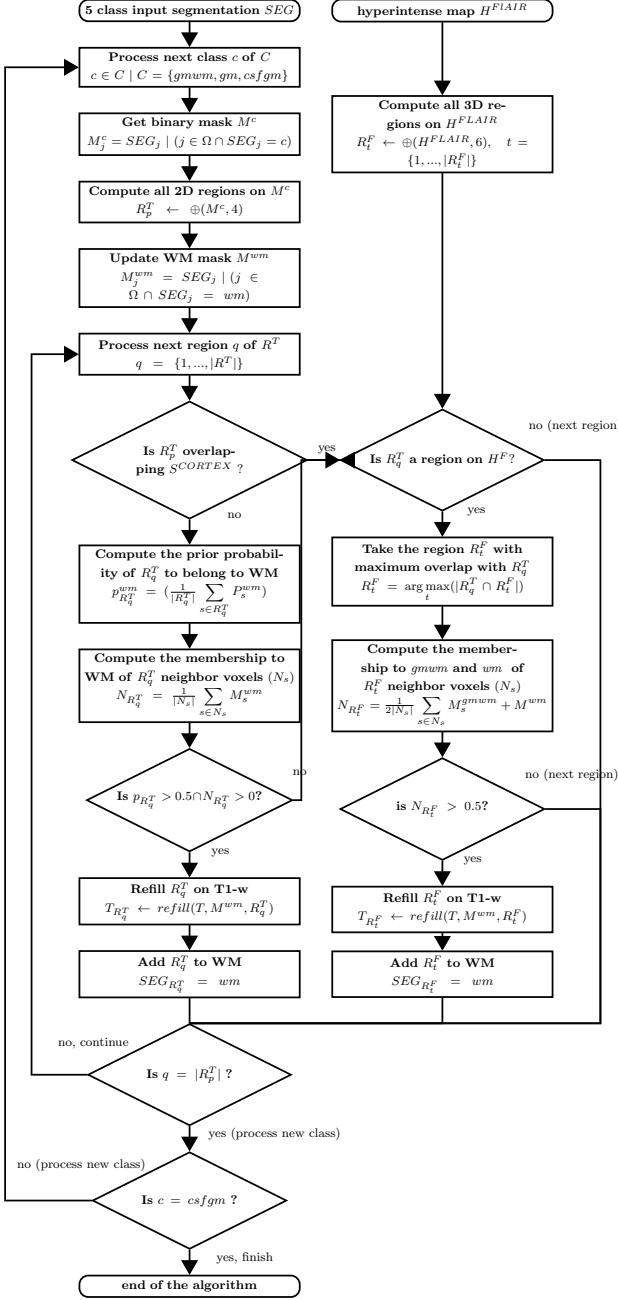


Figure 2: Proposed algorithm to estimate and refill outlier candidate regions into T1-w. The algorithm takes the 5 a class T1-w segmentation and the hyper-intensity map H^{FLAIR} if available as inputs. Connected regions of voxels with similar intensity are filtered based on their spatial location probability on tissue and morphological prior atlases. Selected regions are then refilled on the original T1-w image.

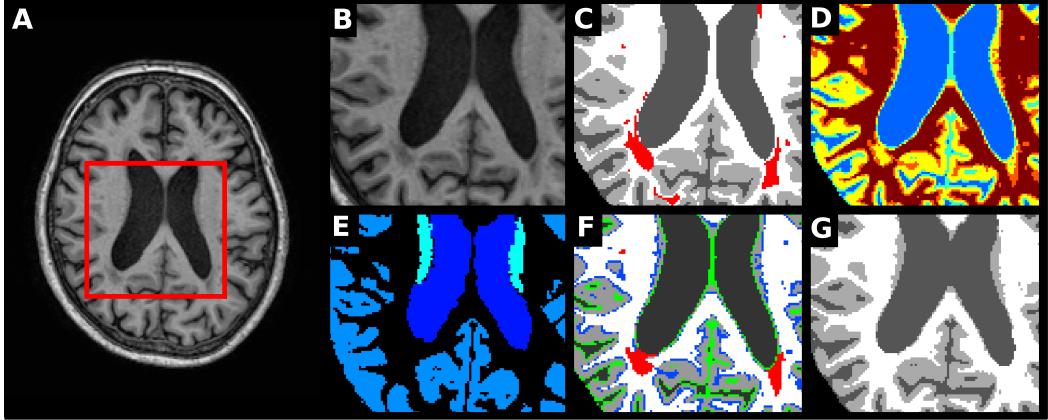


Figure 3: Partial volume assignment during tissue segmentation (Sec. 2.5). A) Original T1-w image. B) Detailed view of the T1-w image. C) Tissue ground-truth with WM lesions highlighted in red. D) Initial 5 class tissue segmentation where *csf*, *csfgm*, *gm*, *gmwm* and *wm* tissues are depicted in blue, cyan, yellow, orange and red, respectively (see Sec. 2.3). E) Morphological brain tissue atlas with parcellated GM regions and ventricles. F) Partial volume *csfgm* and *gmwm* regions (depicted in green and blue, respectively). Previously estimated lesion candidates are depicted in red (see Sec. 2.4). 2D regions with half of their voxels overlapping with the morphological atlas are reassigned to the correspondent tissue. The rest of voxels are reassigned to the neighboring pure class with the most similar signal intensity. G) Final tissue segmentation with partial tissue volume maps re-estimated as CSF (dark gray), GM (light gray) and WM (white).

3. Experiments

3.1. MRBrainS database

3.1.1. Data

The public available MRBrainS 2013 database⁵ consisted of 20 scans with varying degree of brain atrophy and white matter lesions. These scans were acquired on a 3.0T Philips Achieva MR scanner at the University Medical Center Utrecht (The Netherlands) with the following sequences: 3D T1-w (TR: 7.9ms, TE: 4.5ms), T1-Inverse Recovery (TR: 4416ms, TE: 15ms, and TI: 400ms), and T2-weighted/FLAIR (TR: 11000ms, TE: 125ms, and TI: 2800 ms). Each of the scans was co-registered (Klein et al., 2010) and intensity-corrected (Ashburner and Friston, 2005) before releasing the data. T1, T1-IR, and T2/FLAIR voxel size was ($0.96 \times 0.96 \times 3.00 \text{ mm}^3$) after registration (Vincken et al., 2015).

Three experts manually delineated each of the 20 scans into CSF, GM and WM and these annotations were used as the reference standard for the evaluation framework (Vincken et al., 2015). Extended manual annotation containing various brain structures and white matter lesions for 5 scans were provided for training while the remaining 15 scans were blind and had to be skull-stripped and segmented into CSF, GM and WM by participating teams.

3.1.2. Evaluation:

Segmentation results had to be submitted online for external evaluation based on the following scores for the CSF, GM and WM tissues (c):

- Dice similarity coefficient (DSC_c) (Dice, 1945) between the manual tissue segmentation (GT_C) and the computed segmentation (SEG_C) masks:

$$DSC_c = \frac{2 |SEG_c \cap GT_c|}{|SEG_c| + |GT_c|} \times 100 \quad (7)$$

⁵available for download at: <http://mrbrains13.isi.uu.nl/>

- The modified Hausdorff distance (95th percentile) (Huttenlocher et al., 1993) between the manual tissue segmentation (GT_C) points p' and the computed segmentation points p in (SEG_C) masks:

$$h_e^{95} = \max_{p \in SEG_c} \min_{p' \in GT_c} |p - p'| \quad (8)$$

- The absolute difference in tissue volume (AVD_c) between manual tissue segmentation (GT_C) and the computed segmentation (SEG_C) masks:

$$AVD_c = \left\| \frac{|SEG_c| - |GT_c|}{|GT_c|} \right\| \quad (9)$$

In order to evaluate the performance of our method, we submitted two different segmentation sets using either only the T1-w sequences or both T1-w and FLAIR images. We validated the performance of our approach comparing the obtained scores with other submitted segmentation pipelines.

3.1.3. Parameter settings

Skull stripping of input images was performed using a similar approach to other methods participating in the challenge (Jog et al., 2013; Oproeck et al., 2013; Rajchl et al., 2015). The 5 training images were non-rigidly registered to the image space of each of the T1-w (Modat et al., 2010), and the brainmask was generated by simple voting of the registered masks. Afterwards, each mask was refined on the T1-IR image by thresholding hyperintense voxels.

All the parameters of our tissue segmentation method were set to default values ($q = 2$, $\gamma = 0.1$, $w = 1$). The β parameter was computed automatically as described in Section 2.3. The α parameter that scaled the minimum signal intensity on the H^F mask was set to $\alpha = 3$.

3.2. MS database

3.2.1. Data

This non-public database of images was composed by 24 images of clinically isolated syndrome (CIS) patients acquired on a 3T Siemens MR scanner (Trio Tim, Siemens, Germany) with a 12-channel phased-array head coil (data from Hospital Vall D’Hebron, Barcelona, Spain). The following pulse sequences were obtained: 1) transverse proton density and T2-weighted fast spin-echo (TR=2500 ms, TE=16-91 ms, voxel size=0.78×0.78×3mm³); 2) transverse fast T2-FLAIR (TR=9000ms, TE=93ms, TI=2500ms, flip angle=120°, voxel size=0.49×0.49×3mm³); and 3) sagittal 3D T1 magnetization prepared rapid gradient-echo (MPRAGE) (TR=2300 ms, TE=2 ms; flip angle=9°; voxel size=1×1×1.2mm³). For each scan, T1-w and FLAIR images were first skull-stripped using BET (Smith, 2002) and then intensity-corrected using the N3 method (Sled et al., 1998). Finally, FLAIR images were co-registered into the T1-w space and then re-aligned into the MNI space using SPM12 co-registration tools with the normalized mutual information as objective function and trilinear interpolation with no wrapping (Ashburner and Friston, 2005). White matter lesion masks were semi-automatically delineated from FLAIR using JIM software⁶ by an expert radiologist of the hospital center. Mean lesion volume was 4.30 ± 4.84 ml (range 0.1-18.3 ml).

⁶Xinapse Systems, <http://www.xinapse.com/home.php>

3.2.2. Evaluation

Manual expert annotations of tissues were not available for this database. As validated in previous MS studies (Battaglini et al., 2012; Valverde et al., 2015b,c), WM lesions on original T1-w scans were first refilled with signal intensities similar to normal-appearing WM using the SLF lesion filling method (Valverde et al., 2014). Then, both the original and the refilled images were segmented into CSF, GM and WM tissues using our proposed approach. The performance of our tissue segmentation method was evaluated by computing the absolute difference in tissue volume (AVD_c) between the images segmented containing lesions and the same images where WM lesions were refilled before tissue segmentation:

$$AVD_c = \left\| \frac{|SEG_c| - |GT_c^{full}|}{|GT_c^{full}|} \right\| \times 100 \quad (10)$$

where SEG_c refers to the output segmentation masks of the images segmented containing lesions, and GT_c^{full} refers to the output tissue segmentation masks of images where lesions were filled before segmentation and considered as ground-truth.

Several works (DellOglio et al., 2014; Valverde et al., 2015c) have already shown that part of the actual error in tissue segmentation may be partially masked by opposite directions in the differences in total and normal-appearing tissue. In order to add an additional measure estimator of the actual error in tissue segmentation that may not bias by these differences, we also compared for each tissue the percentage of mis-classified voxels PMC_c between the original SEG_c and the expert filled GT_c^{full} masks:

$$PMC_c = \frac{|\overline{SEG_c} \cap GT_c^{full}|}{|GT_c^{full}|} \times 100 \quad (11)$$

In order to analyze the benefits of using the FLAIR image into the proposed approach, we evaluated the performance of our method when using only the T1-w image and when using both T1-w and FLAIR. Furthermore, we validated it with two other automated pipelines widely used in brain tissue segmentation, such as FAST (Zhang et al., 2001) (version FSL 5.0) and SPM (Ashburner and Friston, 2005) (version SPM12 rev 6225), using either original images or after estimating lesions using the automated approach SLS proposed by Roura et al. (2015). On images where lesions were automatically segmented, estimated lesion masks were then filled with the same SLF method (Valverde et al., 2014) before tissue segmentation. Similarly to our approach, we considered the tissue segmentation masks of the expert refilled T1-w images segmented with FAST and SPM12 as the ground-truth for each method. Table 1 summarizes each of the evaluated pipelines and the corresponding process followed to segment the MS images.

3.2.3. Parameter settings

The BET skull-stripped process was optimized as proposed by Popescu et al. (2012) without removing CSF from the brainmask. N3 was run with optimized parameters by reducing the smoothing distance parameter to 30-50 mm (Boyes et al., 2008; Zheng et al., 2009).

The SLF lesion filling method was run with default parameters in all experiments. In the FAST and SPM12 images where we estimated lesion masks automatically, the lesion segmentation method SLS was optimized for 3.0T data identically as shown in Roura et al. (2015).

All the parameters of our proposed method were fixed to default values ($q = 2, \gamma = 0.1, w = 1$) as done in MRBrainS13 database. The β parameter was computed automatically.

Table 1: Summary of evaluated pipelines and processes used on MS data. On pipelines *FAST only T1* and *SPM12 only T1* images were segmented containing lesions without prior automated lesion segmentation. On pipelines *FAST + SLS* and *SPM12 + SLS*, WM lesions were automatically segmented using the SLS approach (Roura et al., 2015) and estimated lesion masks were afterwards lesion filled using the SLF method (Valverde et al., 2014). On our proposed pipeline using *MSSEG only T1* and *MSSEG T1 + FLAIR* lesion segmentation and filling was part of the same segmentation method. Manual lesion annotations were used to refill T1-w images on pipelines *FAST GT*, *SPM12 GT* and *MSSEG GT* before segmenting the images using *FAST*, *SPM12* and *MSSEG*, respectively. AVD_c and PMC_c scores were then computed between pipelines 1 vs 3, 2 vs 3, 4 vs 6, 5 vs 6, 7 vs 9 and 8 vs 9.

Pipeline	Modality	Lesion seg.	Lesion filling	Tissue seg.
1. FAST only T1	T1	<i>none</i>	<i>none</i>	FAST
2. FAST + SLS	T1, FLAIR	SLS (FLAIR)	SLF	FAST
3. FAST GT	T1	<i>expert manual</i>	SLF	FAST
4. SPM12 only T1	T1	<i>none</i>	<i>none</i>	SPM12
5. SPM12 + SLS	T1, FLAIR	SLS (FLAIR)	SLF	SPM12
6. SPM12 GT	T1	<i>expert manual</i>	SLF	SPM12
7. MSSEG only T1	T1	<i>internal</i>	<i>internal</i>	MSSEG
8. MSSEG T1 + FLAIR	T1, FLAIR	<i>internal</i>	<i>internal</i>	MSSEG
9. MSSEG GT	T1	<i>expert manual</i>	SLF	MSSEG

The α parameter that scaled the minimum signal intensity on the H^F mask was set again to $\alpha = 3$.

3.3. Statistical significance

The statistical significance of the performance between methods was computed by running a series of permutation tests (Menke and Martinez, 2004; Klein et al., 2009; Diez et al., 2014) between the differences in the scores obtained by each method. These tests permitted to analyze the fraction of times that a particular method with the lowest score was significantly better than the rest of methods with $p\text{-value} \leq 0.05$. Methods were then ranked in three different levels according to the difference between the mean score of the best method $\mu_o \pm \sigma_o$ and the distance with respect to the mean scores of the rest of methods. Hence, Rank 1 contained methods with mean scores $(\mu_o - \sigma_o, \mu_o]$, Rank 2 contained those with mean scores $(\mu_o - 2\sigma_o, \mu_o - \sigma_o]$ and Rank 3 those in the interval $(\mu_o - 3\sigma_o, \mu_o - 2\sigma_o]$ (Klein et al., 2009; Diez et al., 2014; Valverde et al., 2015a). For all the run tests, we set the number of comparisons between each pair of methods to $N = 1000$.

3.4. Implementation details

The proposed pipeline was entirely developed in MATLAB (v2014a, The Mathworks Inc, US), except for the registration process that was run using the available NiftyReg package (Ourselin et al., 2002; Modat et al., 2010). The method was configured to run either in CPU or GPU. Experiments were carried out on a GNU/Linux machine with a single Intel core i7 processor at 3.4 Ghz (Intel Corp, US), and a NVIDIA K40 with 12GB of RAM (NVIDIA, US). The average execution time for the proposed method including registration and tissue segmentation was 8 minutes running on the CPU core. Execution time on the GPU was approximately 2 minutes, reducing four times the execution time on the CPU processor.

4. Results

4.1. MRBrainS13 dataset

Table ?? shows the obtained mean DSC_c , h_c^{95} and AVD_c scores for our proposed method. We compare the obtained scores with other non-supervised strategies that also participated in the challenge such as FAST, SPM12, or VBM12⁷, and also with respect to the best ranked

⁷<http://www.neuro.uni-jena.de/>

method proposed by Stollenga and Byeon (2015). The overall rank of methods also included the combined brain (GM + WM) and intracranial (CSF + GM + WM) volumes, which are not shown in the table for simplicity⁸. At the time of submitting our results on the online application, our proposed approach using both T1 and FLAIR (*MSSEG T1+FLAIR*) was the best unsupervised method of the challenge (*7th* position overall 31 participants), and its accuracy was also very competitive in comparison with several supervised methods that were explicitly trained for the challenge. When only using the T1-w modality (*MSSEG only T1*), the method was ranked in the *10th* position, but still clearly over-performed *FAST* (*21th* position), and *SPM12* using FLAIR+T1 (*17th* position), the T1-IR modality (*18th* position), or the T1-w modality (*20th* position).

Figure 4 illustrates the different steps performed by our approach. After registering the probabilistic atlases into the subject space (Fig.4 panels E to H), tissue was estimated from the T1-w into 5 different classes (Fig.4 panel I). Then, WM outliers were estimated on the T1-w image by analyzing all the regions not initially segmented as WM with a high probability to pertain to WM based on spatial local probability and prior tissue information (red regions depicted in Fig.4 panel J). If FLAIR was also provided, lesion candidates were also analyzed based on their signal intensity on the FLAIR image and their spatial local probability to pertain to WM (green regions depicted in Fig.4 panel J). Afterwards, lesion candidate regions were refilled into the T1-w image with signal intensities similar to the WM, the refilled T1-w image was re-estimated again and CSFGM and GMWM volumes were reassigned into the three main classes (Fig.4 panels K and L).

4.2. MS data

Table 2 (a) depicts the mean % of absolute differences in CSF, GM and WM volume (AVD_c) for each of the evaluated methods after segmenting the 24 3T images. Table 2 (b) shows the mean % of mis-classified CSF, GM and WM voxels (PMC_c) for each of the evaluated methods.

Table 2: Mean % of absolute difference in CSF, GM and WM volume between the 24 3T tissue masks where expert annotations were refilled before segmentation and the same images segmented including white matter lesions. For each method, reported values are the mean and standard deviation $\mu \pm \sigma$ for the (a) AVD_c and (b) PMC_c scores obtained along the entire database.

(a) Differences in AVD_c

Method	Dif CSF (%)	Dif GM (%)	Dif WM (%)
FAST only T1	0.07 ± 0.13	0.33 ± 0.45	0.42 ± 0.56
FAST + SLS	0.04 ± 0.07	0.08 ± 0.12	0.11 ± 0.16
SPM12 only T1	0.31 ± 0.46	0.27 ± 0.45	0.56 ± 0.69
SPM12 + SLS	0.22 ± 0.22	0.13 ± 0.23	0.20 ± 0.32
MSSEG only T1	0.13 ± 0.20	0.21 ± 0.26	0.42 ± 0.54
MSSEG T1+FLAIR	0.04 ± 0.05	0.06 ± 0.05	0.13 ± 0.11

(b) Differences in PMC_c

Method	CSF (%)	GM (%)	WM (%)
FAST only T1	0.08 ± 0.11	0.09 ± 0.12	0.53 ± 0.69
FAST + SLS	0.06 ± 0.06	0.14 ± 0.16	0.25 ± 0.30
SPM12 only T1	0.16 ± 0.32	0.25 ± 0.33	0.73 ± 0.86
SPM12 + SLS	0.22 ± 0.31	0.24 ± 0.29	0.41 ± 0.43
MSSEG only T1	0.02 ± 0.03	0.03 ± 0.05	0.46 ± 0.58
MSSEG T1+FLAIR	0.04 ± 0.04	0.14 ± 0.13	0.27 ± 0.29

⁸Overall ranking of methods for all the measurements can be consulted at <http://mrbrains13.isi.uu.nl/results.php>

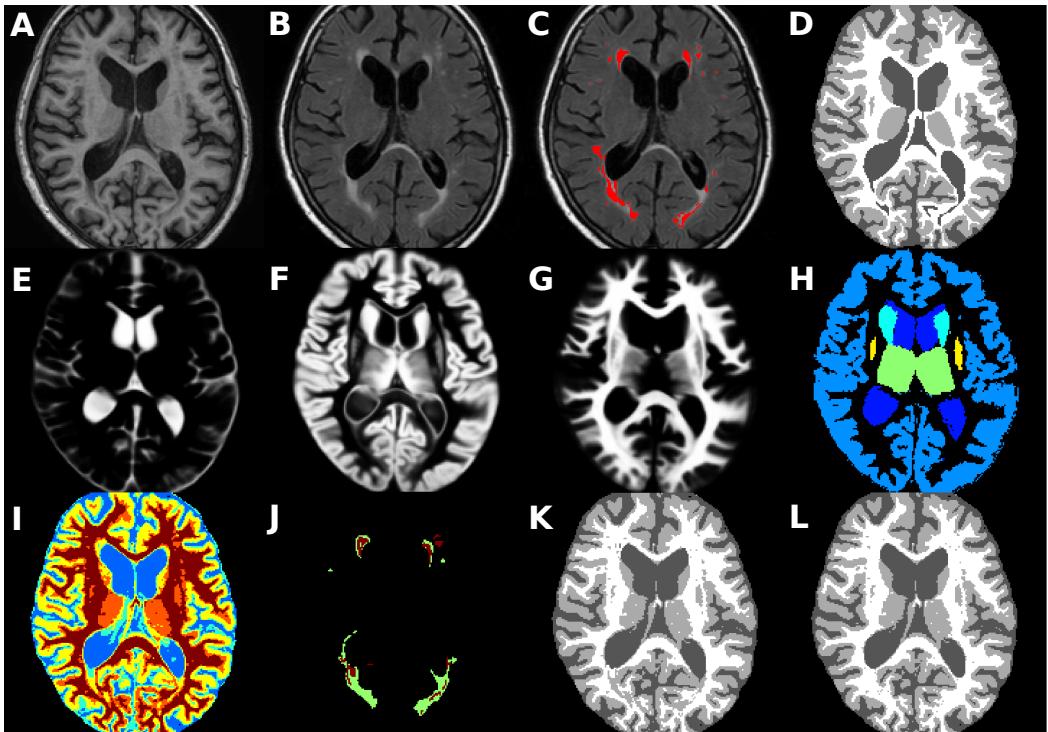


Figure 4: Automated tissue segmentation of the *MSSEG* method on the *second subject* of the training set of the MRBrainS13 database. A) Original T1-w image. B) Original FLAIR image. C) FLAIR image with manual annotated WM lesions depicted in red. D) Provided ground-truth for training purposes. Registered CSF, GM and WM prior atlas to the subject space (E, F and G, respectively). H) Morphological brain structural atlas registered to the subject space. I) First partial volume segmentation with *csf* depicted in blue, *csfgm* in cyan, *gm* in yellow, *gmwm* in orange and *wm* in red. J) Obtained WM outliers extracted from either T1-w (depicted in red) and FLAIR (depicted in green). K) Final tissue segmentation using only the T1-w image, with *CSF* depicted in dark gray, *GM* in light gray and *WM* in white. L) Final tissue segmentation when using both T1 and the FLAIR images.

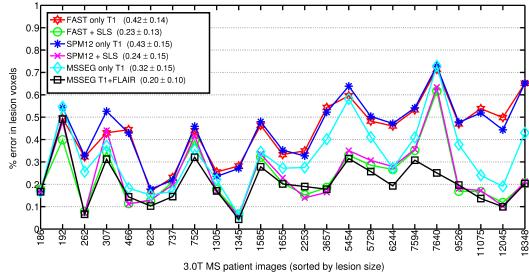


Figure 5: % of absolute difference in WM lesion volume for each of the evaluated pipelines on the 3T MS database. Figure legends also depict mean and standard deviation $\mu \pm \sigma$ for the entire set of images. Images are sorted by lesion size (number of lesion voxels).

Methods were ranked based on their scores after running the permutation tests. Table 3 shows the rank of each evaluated method for the AVD_c (Table 3 (a)) and PMC_C scores (Table 3 (b)), respectively. The differences in tissue volume were the lowest when methods used also the FLAIR image to estimate the WM outliers. Our proposed approach reported competitive results and again was ranked in the first rank of methods for all tissues. Methods using only T1 yielded a significant higher difference in GM and WM volume and were ranked in the second and third group. The proposed approach using only T1-w was significantly better than the rest of methods in terms of the % of miss-classified GM and WM voxels, but its performance was worst for WM and was ranked in the second group of methods for WM. As shown in Table 3 (a), the % of miss-classified WM voxels was significantly lower in *FAST+SLS* and our proposed method *MSSEG T1+FLAIR* when compared with the rest of evaluated pipelines.

4.2.1. WM outlier rejection

Finally, we evaluated the performance of the proposed WM outlier rejection algorithm with respect to the rest of pipelines. A traditional comparison respect of the number of true-positive and false-positive WM lesion voxels of each pipeline is not appropriate here, given that our approach only processed WM lesion candidates that were not initially classified as WM. In contrast, the error in the expected WM lesion volume segmented can be a good indicator of the performance of each method segmenting WM lesions as WM.⁹ Figure 5 shows the % of absolute difference in WM lesion volume between each of the evaluated pipelines and their correspondent GT_c^{fill} images. As expected, methods yielded the lowest differences in WM lesion volume when used also the FLAIR modality to estimate WM lesions. *MSSEG+FLAIR* showed the lowest differences in WM lesion volume of all evaluated methods.

5. Discussion

In this paper we have presented a new automated brain tissue segmentation pipeline for MS patient images that combines multi-channel intensities, anatomical and morphological prior maps at different levels to estimate brain tissue in the presence of WM lesions. The current method integrates a WM outlier estimation and refilling algorithm which is applied intermediately in order to reduce the effect of WM lesions on tissue segmentation. As shown by the presented results, the proposed technique yields competitive and consistent results in both general and MS specific databases without parameter tweaking. Furthermore, although

⁹Aquest punt s'ha de revisar. cal parlar-ne.

Table 3: Permutation tests results for evaluated methods on the 3T MS database. (a) Final rank based on the absolute % difference in CSF, GM and WM volume between methods. (b) Final rank based on % of miss-classified CSF, GM and WM voxels between methods. Reported values are mean and standard deviation (μ_o, σ_o) of the fraction of times when each method produces significant p-values ($p \leq 0.05$). Positive values indicate that in average, the method out-performed the other methods in pair-wise significant tests. Negative values indicate the contrary. Rank 1: $(\mu_o - \sigma_o, \mu_o]$, Rank 2: $(\mu_o - 2\sigma_o, \mu_o - \sigma]$, Rank 3 $(\mu_o - 3\sigma_o, \mu_o - 2\sigma_o]$. All permutation tests were run with 1000 random iterations.

(a) Evaluated methods ranked by the absolute % of CSF, GM and WM volume of 3T data.						
Rank	Method (CSF)	$\mu \pm \sigma$	Method (GM)	$\mu \pm \sigma$	Method (WM)	$\mu \pm \sigma$
Rank 1	MSSEG T1+FLAIR	0.5 ± 0.55	MSSEG T1+FLAIR	0.5 ± 0.55	FAST + SLS	0.67 ± 0.52
	FAST only T1	0.5 ± 0.55	FAST + SLS	0.5 ± 0.55	MSSEG T1+FLAIR	0.5 ± 0.55
	FAST + SLS	0.5 ± 0.55	SPM12 + SLS	0.33 ± 0.52		
Rank 2	MSSEG only T1	-0.43 ± 0.64	MSSEG only T1	-0.17 ± 0.75	SPM12 + SLS	0.14 ± 0.72
	SPM12 + SLS	-0.5 ± 0.55	SPM12 only T1	-0.5 ± 0.55	FAST only T1	-0.24 ± 0.62
	SPM12 only T1	-0.57 ± 0.5				
Rank 3		FAST only T1		-0.67 ± 0.52	MSSEG only T1	-0.47 ± 0.52
					SPM12 only T1	-0.59 ± 0.49

(b) Evaluated methods ranked by the absolute % of miss-classified CSF, GM and WM of 3T data.						
Rank	Method (CSF)	$\mu \pm \sigma$	Method (GM)	$\mu \pm \sigma$	Method (WM)	$\mu \pm \sigma$
Rank 1	MSSEG only T1	0.83 ± 0.41	MSSEG only T1	0.83 ± 0.41	MSSEG T1+FLAIR	0.67 ± 0.52
	MSSEG T1+FLAIR	0.44 ± 0.8	FAST only T1	0.5 ± 0.84	FAST + SLS	0.67 ± 0.52
Rank 2					SPM12 + SLS	-0.17 ± 0.75
					MSSEG only T1	-0.17 ± 0.75
					FAST only T1	-0.17 ± 0.75
Rank 3	FAST + SLS	0 ± 0.89	FAST + SLS	-0.17 ± 0.75	SPM12 only T1	-0.83 ± 0.41
	SPM12 only T1	-0.27 ± 0.43	MSSEG T1+FLAIR	-0.17 ± 0.75		
	FAST only T1	-0.33 ± 0.82	SPM12 only T1	-0.33 ± 0.52		
	SPM12 + SLS	-0.67 ± 0.52	SPM12 + SLS	-0.66 ± 0.51		

we did not analyze explicitly the execution times of each of the evaluated algorithms, the proposed method takes advantage of new affordable processors such as GPUs that reduce up to four times the execution time to register and segment tissue when compared to general purpose CPUs.

The MRBrainS challenge permitted to evaluate the efficacy of our method and to validate it with other state-of-the-art tissue segmentation methods. Although the challenge was not focused on the MS disease, methods were evaluated by comparing them with respect manual expert annotations of tissues and WM lesions, which provided a quantitative measure of the accuracy of the method. The overall results showed that supervised methods obtained the best results of the challenge, taking advantage of the inherent capability to fit the database characteristics. At the time of writing this paper, *MSSSEG T1+FLAIR* was ranked in the 7th position out of 31 participants, being the best non-supervised strategy followed by the VBM12 approach. As shown by the differences at each of the obtained scores, the FLAIR modality appeared to be useful to improve the accuracy of the method when compared with *MSSSEG T1*, which was ranked 10th. The performance of *MSSEG* was superior in all tissues when compared to general purpose methods such as *FAST* (ranked 21th) and *SPM12* (best ranked 17th), even if those used both image modalities. However, final ranking of methods should be taken with care, given the differences in the skull-stripping processes between methods. Differences in the boundaries of the estimated skull masks may be behind the remarkable differences in CSF between methods, altering also the intra-cranial cavity measurements and consequently the overall score of each of the methods.

In MS data, the performance of our method was similar or better to the best pipeline incorporating an state-of-the-art method for lesion segmentation and filling, validating the overall capability of the proposed method to reduce the effects of WM lesions on tissue seg-

mentation. *MSSEG T1+FLAIR* and *FAST+SLS* were ranked in the first group of methods with error differences in tissue volume below 0.15% in all tissues. Pipelines using only T1-w showed a similar or lower % of miss-classified CSF and GM voxels than those using both FLAIR and T1-w. In contrast, the % of miss-classified WM voxels (Table 3 (a)) and the differences in reassigned lesion volume (Fig. 5) were significantly higher on the former, showing that these methods tended to overestimate GM and underestimate WM by the effect of WM lesions. In this aspect, our results are consistent with previous studies also analyzing the effects of WM lesions on tissue segmentation (Battaglini et al., 2012; Gelineau-Morel et al., 2012; Valverde et al., 2015b,c).

Differences in the AVD_c between *MSSEG T1+FLAIR* and *MSSEG only T1* on the MRBrains13 data were similar than those reported in MS data, showing that in general the inclusion of the FLAIR modality reduced the overall error in tissue volume on all the analyzed databases. On MS data, the % of miss-classified CSF and GM voxels was significantly lower on the *MSSEG only T1*, but significantly higher in WM, evidencing that *MSSEG only T1* tended to overestimate WM, while the error in *MSSEG T1+FLAIR* was similar in both GM and WM. In addition, the presented results show that the % difference in total WM and lesion volume was significantly lower on the *MSSEG T1+FLAIR* in comparison with *MSSEG only T1*. Hence, we would recommend to use both T1-w and FLAIR modalities when possible. However, the accuracy of the *MSSEG only T1* pipeline was still superior than *FAST* and *SPM12* when compared with ground-truth annotations of the MRBrainS13 database. This suggests that at least with the available data, the improvement in tissue segmentation was not only caused by the addition of the FLAIR modality, but also by the combination of intensity, anatomical and morphological priors.

This study however endorses some limitations. The lack of a database consisting of MS images with manual annotations of tissue, limits our analysis to the differences in tissue volume with respect to images where expert lesion annotations were lesion filled before tissue segmentation. However, the previous analysis has been shown in previous studies to be effective to evaluate the effects of WM lesions on tissue segmentation (Battaglini et al., 2012; Valverde et al., 2015b,c). Furthermore, the mean lesion sizes of the MS cohorts do not allow to investigate better the performance of the proposed method in the presence of images with higher lesion load. As a future work, we believe that an additional study on MS with manual tissue annotated masks and higher lesion load would be helpful not only to analyze the benefits of the proposed algorithm in MS images, but also to investigate the benefits of adding other image channels such as T2 or PD. Furthermore, although the method was designed for cross-sectional data, we are sensible to the fact that the current approach may be benefited by the possibility to evaluate longitudinal changes in tissue volume.

6. Conclusion

In this paper, we have proposed the Multiple Sclerosis SEGmentation pipeline (*MSSEG*), a new MRI brain tissue segmentation method designed to deal with MS patient images containing lesions. Our proposed approach incorporates robust partial volume tissue segmentation with outlier rejection and filling, combining intensity, probabilistic and morphological prior maps in a novel-way. When combining T1-w and FLAIR modalities, our method have shown very competitive results on the MRBrainS13 database, ranked on the 7th position out of 31 participant strategies and being the best non-supervised approach so far. In MS data, differences in tissue volume were lower or similar to the best available pipeline composed of *FAST* and a state-of-the-art method for lesion segmentation and filling. In all the experiments the inclusion of the FLAIR modality into the proposed method reduced the effect of

WM lesions on tissue segmentation, which suggests that this modality should be used when available. In conclusion, our results show that at least with the presented data, *MSSEG* improves the measurement of brain tissue volume on images containing WM lesions. Hence, we strongly believe that the neuro-image community can be benefited by its use in future settings.

Acknowledgements:

S. Valverde holds a FI-DGR2013 grant from the Generalitat de Catalunya. E. Roura holds a BR-UdG2013 grant. This work has been partially supported by "La Fundació la Marató de TV3", by Retos de Investigación TIN2014-55710-R, and by the MPC UdG 2016/022 grant. The authors gratefully acknowledge the support of NVIDIA Corporation with the donation of the Tesla K40 GPU used for this research.

References

- Ashburner, J. and Friston, K. J. (2005). Unified segmentation. *NeuroImage*, 26:839–851.
- Battaglini, M., Jenkinson, M., Stefano, N. D., and De Stefano, N. (2012). Evaluating and reducing the impact of white matter lesions on brain volume measurements. *Human Brain Mapping*, 33(9):2062–71.
- Boyes, R. G., Gunter, J. L., Frost, C., Janke, A. L., Yeatman, T., Hill, D. L. G., Bernstein, M. A., Thompson, P. M., Weiner, M. W., Schuff, N., Alexander, G. E., Killiany, R. J., DeCarli, C., Jack, C. R., and Fox, N. C. (2008). Intensity non-uniformity correction using N3 on 3-T scanners with multichannel phased array coils. *NeuroImage*, 39:1752–1762.
- Cabezas, M., Oliver, A., Lladó, X., Freixenet, J., and Bach Cuadra, M. (2011). A review of atlas-based segmentation for magnetic resonance brain images. *Computer Methods and Programs in Biomedicine*, 104(3):e158–e177.
- Ceccarelli, A., Jackson, J. S., Tauhid, S., Arora, A., Gorky, J., Dell’Oglio, E., Bakshi, A., Chitnis, T., Khouri, S. J., Weiner, H. L., Guttmann, C. R. G., Bakshi, R., and Neema, M. (2012). The impact of lesion in-painting and registration methods on voxel-based morphometry in detecting regional cerebral gray matter atrophy in multiple sclerosis. *American Journal of Neuroradiology*, 33(8):1579–85.
- Chard, D. T., Jackson, J. S., Miller, D. H., and Wheeler-Kingshott, C. A. M. (2010). Reducing the impact of white matter lesions on automated measures of brain gray and white matter volumes. *Journal of Magnetic Resonance Imaging*, 32:223–228.
- Datta, S. and Narayana, P. a. (2013). A comprehensive approach to the segmentation of multichannel three-dimensional MR brain images in multiple sclerosis. *NeuroImage: Clinical*, 2:184–96.
- De Boer, R., Vrooman, H. a., van der Lijn, F., Vernooij, M. W., Ikram, M. A., van der Lugt, A., Breteler, M. M. B., and Niessen, W. J. (2009). White matter lesion extension to automatic brain tissue segmentation on MRI. *NeuroImage*, 45(4):1151–61.
- de Bresser, J., Portegies, M. P., Leemans, A., Biessels, G. J., Kappelle, L. J., Viergever, M. a., Bresser, J. D., and Jan, G. (2011). A comparison of MR based segmentation methods for measuring brain atrophy progression. *NeuroImage*, 54(2):760–8.

- DellOglio, E., Ceccarelli, A., Glanz, B. I., Healy, B. C., Tauhid, S., Arora, A., Saravanan, N., Bruha, M. J., Vartanian, A. V., Dupuy, S. L., Benedict, R. H., Bakshi, R., and Neema, M. (2014). Quantification of Global Cerebral Atrophy in Multiple Sclerosis from 3T MRI Using SPM: The Role of Misclassification Errors. *Journal of Neuroimaging*, 25(2):191–199.
- Deshpande, H., Maurel, P., and Barillot, C. (2015). Classification of Multiple Sclerosis Lesions using Adaptive Dictionary Learning. *Computerized Medical Imaging and Graphics*.
- Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302.
- Diez, Y., Oliver, A., Cabezas, M., Valverde, S., Martí, R., Vilanova, J. C., Ramió-Torrentà, L., Rovira, À., and Lladó, X. (2014). Intensity based methods for brain MRI longitudinal registration. A study on multiple sclerosis patients. *Neuroinformatics*, 12(3):365–379.
- García-Lorenzo, D., Francis, S., Narayanan, S., Arnold, D. L., and Collins, D. L. (2013). Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Medical Image Analysis*, 17(1):1–18.
- Gelineau-Morel, R., Tomassini, V., Jenkinson, M., Johansen-berg, H., Matthews, P. M., and Palace, J. (2012). The Effect of Hypointense White Matter Lesions on Automated Gray Matter Segmentation in Multiple Sclerosis. *Human Brain Mapping*, 2814(June 2011):2802–2814.
- Geremia, E., Clatz, O., Menze, B. H., Konukoglu, E., Criminisi, A., and Ayache, N. (2011). Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images. *NeuroImage*, 57(2):378–390.
- Giorgio, A. and De Stefano, N. (2013). Clinical use of brain volumetry. *Journal of Magnetic Resonance Imaging*, 37(1):1–14.
- Guizard, N., Coupé, P., Fonov, V. S., Manjón, J. V., Arnold, D. L., and Collins, D. L. (2015). Rotation-invariant multi-contrast non-local means for MS lesion segmentation. *NeuroImage: Clinical*, 8:376–389.
- Harmouche, R., Subbanna, N. K., Collins, D. L., Arnold, D. L., and Arbel, T. (2015). Based on Modeling Regional Intensity Variability and Local Neighborhood Information. *IEEE Transactions on Biomedical Engineering*, 62(5):1281–1292.
- Huttenlocher, D., Klanderman, G., and Ruckridge, W. (1993). Comparing images using the hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850–863.
- Immerkær, J. (1996). Fast Noise Variance Estimation. *Computer Vision and Image Understanding*, 64(2):300–302.
- Jain, S., Sima, D. M., Ribbens, A., Cambron, M., Maertens, A., Van Hecke, W., De Mey, J., Barkhof, F., Steenwijk, M. D., Daams, M., Maes, F., Van Huffel, S., Vrenken, H., and Smeets, D. (2015). Automatic segmentation and volumetry of multiple sclerosis brain lesions from MR images. *NeuroImage: Clinical*, 8:367–375.
- Jog, A., Roy, S., Prince, J., and Carass, A. (2013). Mr brain segmentation using decision trees. In *Proceedings of the MICCAI Workshops, Challenge on MR Brain Image Segmentation (MRBrainS 13), The Midas Journal*.

- Klauschen, F., Goldman, A., Barra, V., Meyer-Lindenberg, A., and Lundervold, A. (2009). Evaluation of automated brain MR image segmentation and volumetry methods. *Human brain mapping*, 30(4):1310–27.
- Klein, A., Ardekani, B., Andersson, J., Ashburner, J., Avants, B., Chiang, M., Christensen, G., Collins, D., Gee, J., Hellier, P., and Others (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, 46(3):786–802.
- Klein, S., Staring, M., Murphy, K., Viergever, M. a., and Pluim, J. P. W. (2010). Elastix: A toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205.
- Lladó, X., Oliver, A., Cabezas, M., Freixenet, J., Vilanova, J. C., Quiles, A., Valls, L., Ramió-Torrentà, L., and Rovira, A. (2012). Segmentation of multiple sclerosis lesions in brain MRI: A review of automated approaches. *Information Sciences*, 186(1):164–185.
- Magon, S., Gaetano, L., Chakravarty, M. M., Lerch, J. P., Naegelin, Y., Stippich, C., Kappos, L., Radue, E.-W., and Sprenger, T. (2014). White matter lesion filling improves the accuracy of cortical thickness measurements in multiple sclerosis patients: a longitudinal study. *BMC Neuroscience*, 15(1):106.
- Menke, J. and Martinez, T. (2004). Using permutations instead of student's t distribution for p-values in paired-difference algorithm comparisons. In *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, volume 2, pages 1331–1335 vol.2.
- Modat, M., Ridgway, G. R., Taylor, Z. A., Lehmann, M., Barnes, J., Hawkes, D. J., Fox, N. C., and Ourselin, S. (2010). Fast free-form deformation using graphics processing units. *Computer methods and programs in biomedicine*, 98(3):278–84.
- Nakamura, K. and Fisher, E. (2009). Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients. *NeuroImage*, 44(3):769–76.
- Opbroek, A. V., der Lijn, F. V., and Bruijne, M. D. (2013). Automated brain-tissue segmentation by multi-feature svm classification. In *Proceedings of the MICCAI Workshops, Challenge on MR Brain Image Segmentation (MRBrainS 13), The Midas Journal*.
- Ourselin, S., Stefanescu, R., and Pennec, X. (2002). Robust registration of multi-modal images: Towards real-time clinical applications. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI'02*, pages 140–147. Springer.
- Pérez-Miralles, F., Sastre-Garriga, J., Tintoré, M., Arrambide, G., Nos, C., Perkal, H., Río, J., Edo, M. C., Horga, a., Castilló, J., Auger, C., Huerga, E., Rovira, a., and Montalban, X. (2013). Clinical impact of early brain atrophy in clinically isolated syndromes. *Multiple sclerosis (Hounds Mills, Basingstoke, England)*, 19(14):1878–86.
- Pham, D. L. (2001). Spatial Models for Fuzzy Clustering. *Computer Vision and Image Understanding*, 297:285–297.
- Pohl, K. M., Bouix, S., Nakamura, M., Rohlfing, T., McCarley, R. W., Kikinis, R., Grimson, W. E. L., Shenton, M. E., and Wells, W. M. (2007). A hierarchical algorithm for MR brain image parcellation. *IEEE transactions on medical imaging*, 26(9):1201–1212.

- Popescu, V., Battaglini, M., Hoogstrate, W. S., Verfaillie, S. C. J., Sluimer, I. C., van Schijndel, R. A., van Dijk, B. W., Cover, K. S., Knol, D. L., Jenkinson, M., Barkhof, F., de Stefano, N., Vrenken, H., Barkhof, F., Montalban, X., Fazekas, F., Filippi, M., Frederiksen, J., Kappos, L., Miller, D., Palace, J., Polman, C., Rocca, M., Rovira, A., and Yousry, T. (2012). Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis. *NeuroImage*, 61:1484–1494.
- Popescu, V., Ran, N. C. G., Barkhof, F., Chard, D. T., Wheeler-Kingshott, C., and Vrenken, H. (2014). Accurate GM atrophy quantification in MS using lesion-filling with co-registered 2D lesion masks. *NeuroImage: Clinical*, 4:366–73.
- Rajchl, M., Baxter, J. S., McLeod, a. J., Yuan, J., Qiu, W., Peters, T. M., and Khan, A. R. (2015). Hierarchical max-flow segmentation framework for multi-atlas segmentation with Kohonen self-organizing map based Gaussian mixture modeling. *Medical Image Analysis*, 000:1–12.
- Roura, E., Oliver, A., Cabezas, M., Valverde, S., Pareto, D., Vilanova, J. C., Ramió-Torrentà, L., Rovira, À., and Lladó, X. (2015). A toolbox for multiple sclerosis lesion segmentation. *Neuroradiology*, 57(10):1031–1043.
- Roy, S., He, Q., Sweeney, E., Carass, A., Reich, D. S., Prince, J. L., and Pham, D. L. (2015). Subject-Specific Sparse Dictionary Learning for Atlas-Based Brain MRI Segmentation. *IEEE J. Biomed. Heal. informatics*, 19(5):1598–609.
- Schmidt, P., Gaser, C., Arsic, M., Buck, D., Förtschler, A., Berthele, A., Hoshi, M., Ilg, R., Schmid, V. J., Zimmer, C., Hemmer, B., and Mühlau, M. (2012). An automated tool for detection of FLAIR-hyperintense white-matter lesions in Multiple Sclerosis. *NeuroImage*, 59(4):3774–3783.
- Shiee, N., Bazin, P.-L., Ozturk, A., Reich, D. S., Calabresi, P. A., and Pham, D. L. (2010). A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. *NeuroImage*, 49(2):1524–1535.
- Sled, J. G., Zijdenbos, a. P., and Evans, a. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, 17(1):87–97.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17:143–155.
- Smith, S. M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P., Federico, A., and De Stefano, N. (2002). Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis . *NeuroImage*, 17(1):479–489.
- Sormani, M. P., Arnold, D. L., and De Stefano, N. (2014). Treatment effect on brain atrophy correlates with treatment effect on disability in multiple sclerosis. *Annals of neurology*, 75(1):43–9.
- Stollenga, M. and Byeon, W. (2015). [M] Parallel Multi-Dimensional LSTM, With Application to Fast Biomedical Volumetric Image Segmentation. In *Neural Information Processing Systems Conference*.

- Sweeney, E. M., Shinohara, R. T., Shiee, N., Mateen, F. J., Chudgar, A. a., Cuzzocreo, J. L., Calabresi, P. a., Pham, D. L., Reich, D. S., and Crainiceanu, C. M. (2013). OASIS is Automated Statistical Inference for Segmentation, with applications to multiple sclerosis lesion segmentation in MRI. *NeuroImage: Clinical*, 2:402–13.
- Tomas-Fernandez, X. and Warfield, S. K. (2015). A Model of Population and Subject (MOPS) Intensities with Application to Multiple Sclerosis Lesion Segmentation. *IEEE transactions on medical imaging*, 0062(c):1–15.
- Valverde, S., Oliver, A., Cabezas, M., Roura, E., and Lladó, X. (2015a). Comparison of 10 brain tissue segmentation methods using revisited IBSR annotations. *Journal of Magnetic Resonance Imaging*, 41(1):93–101.
- Valverde, S., Oliver, A., Dez, Y., Cabezas, M., Vilanova, J., Rami-Torrentà, L., Rovira, A., and Lladó, X. (2015b). Evaluating the effects of white matter multiple sclerosis lesions on the volume estimation of 6 brain tissue segmentation methods. *American Journal of Neuroradiology*, 36(6):1109–1115.
- Valverde, S., Oliver, A., and Lladó, X. (2014). A white matter lesion-filling approach to improve brain tissue volume measurements. *NeuroImage: Clinical*, 6:86–92.
- Valverde, S., Oliver, A., Roura, E., Pareto, D., Vilanova, J. C., Ramió-Torrentà, L., Sastre-Garriga, J., Montalban, X., Rovira, A., and Lladó, X. (2015c). Quantifying brain tissue volume in multiple sclerosis with automated lesion segmentation and filling. *NeuroImage: Clinical*, 9:640–647.
- Vincken, K. L., Kuijf, H. J., Breeuwer, M., Bouvy, W. H., Bresser, J. D., Alansary, A., Bruijne, M. D., Carass, A., El-baz, A., Jog, A., Katyal, R., Khan, A. R., Lijn, F. V. D., Mahmood, Q., Mukherjee, R., Opbroek, A. V., Paneri, S., Persson, M., Rajchl, M., Sarikaya, D., Silva, C. A., Vrooman, H. A., Vyas, S., Wang, C., Zhao, L., Biessels, G. J., and Viergever, M. A. (2015). MRBrainS Challenge: Online Evaluation Framework for Brain Image Segmentation in 3T MRI Scans. *Computational Intelligence and Neuroscience*.
- Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20:45–57.
- Zheng, W., Chee, M. W. L., and Zagorodnov, V. (2009). Improvement of brain segmentation accuracy by optimizing non-uniformity correction using N3. *NeuroImage*, 48:73–83.

Chapter 7

Main results and discussion

MRI tissue segmentation techniques are increasingly being used as the standard tools to assess brain tissue volume. However, automated tissue segmentation is still a challenging task in MS, due to tissue abnormalities found in MS image patients such as WM lesions that are known to reduce the accuracy of tissue segmentation methods. When expert manual annotations of WM lesions are available, lesion filling has been shown to be an effective method to reduce the effects of those lesions on tissue segmentation. However, manual annotations are time-consuming and prone to variability between experts, which in combination of the need to analyze quantitatively focal MS lesions in individual and temporal studies has led to the development of a wide number of automated lesion segmentation of MS lesions. Therefore, a solid understanding of the effects of MS lesions on automated pipelines that concatenate processes such as lesion segmentation, lesion filling and tissue segmentation is important.

As said in section 1.3, each of these processes cover part of the necessary knowledge needed to tackle the problem of automated brain tissue segmentation of images containing lesions. This chapter provides a comprehensive discussion of the main results obtained on previous chapters, analyzing each of these necessary processes until the development of a fully automated tissue segmentation method for MS images.

7.1 Effect of WM lesions on tissue segmentation

A wide number of automated tissue segmentation methods have been proposed in the literature so far. In Chapter 2, we evaluated the accuracy of ten approaches using the two public available databases of healthy subjects IBSR18 and IBSR20. With the aim to include a wide set of different segmentation approaches and available tools, the analysis included well-known intensity based algorithms such as ANN, FCM,

and KNN, and also public available toolboxes such as FAST, SPM5, SPM8, PVC, GAMIXTURE, SVPASEG, and FANTASM. Results were presented before and after correcting CSF masks, as it was shown that available annotations ignored sulcal CSF tissue on original masks. When sulcal CSF was corrected, SVPASEG, SPM8 and FAST yielded the highest accuracy on both databases. However, most of the methods were variant to changes in acquisition sequences, intensity inhomogeneities or special attributes of the different databases, which highlighted the fact that brain tissue segmentation problem was still open, because there was not a single method that achieved a significant higher accuracy on all tissues.

Afterwards, six of these methods (ANN, FCM, FANTASM, FAST, SPM5 and SPM8) were evaluated on MS data from different hospital centers and scanners in order to analyze the extent to which tissue volume estimations were affected by changes in WM lesion volume and intensity. Our results showed that SPM8 was the method with the lowest differences in total volume, while FANTASM and again SPM8 were the methods where the incidence of WM lesions was the lowest on normal-appearing tissue. In general, differences in tissue volume were lower on methods combining morphological prior information such as SPM5 and SPM8, or spatial constraints such as FANTASM and FAST. In contrast, these differences were higher on simpler intensity based algorithms that lacked of spatial correction such as FCM and ANN. This fact and the higher performance on healthy data of former methods, stressed the necessity of adding morphological prior information and/or spatial constraints in automated brain tissue segmentation, not only to overcome inherent MRI artifacts but also as an important component to deal with WM lesions.

The main factor in the observed differences in tissue volume across methods was caused by lesion volume. Furthermore, WM lesion voxels tended to be classified as GM on images where the variation between lesion signal intensity and the mean signal intensity of normal-appearing WM was higher, which indicates a direct relationship between the differences in brain tissue volume and changes in lesion load and WM lesion intensity. However, lesion voxels had also a direct effect on the observed differences in GM and WM outside lesion regions. As already commented in Chapter 3, these differences are especially important because they highlight the bias introduced by WM lesions on the estimation of tissue volume that is not pathologically affected. Our analysis showed that if lesion voxels were not considered to compute brain volume, still methods tended to overestimate GM, specially on images with higher lesion load. Observed differences in normal-appearing tissue volume were important, because although lesion voxels could be reassigned to WM after segmentation, if these lesions were present in image segmentation, part of the bias was still present. Furthermore, differences in total tissue volume may be canceled between the errors produced in the same lesion regions and the effect of these voxels in normal-appearing tissue. This fact clearly shows the necessity to process

WM lesion before segmentation.

7.2 Effect of lesion filling in tissue segmentation

In the last years, different techniques have been proposed to reduce the bias introduced by WM lesions on brain tissue volume measurements of MS images, mostly by in-painting WM lesions on T1-w with signal intensities similar to normal-appearing WM. After reviewing the related available literature in Chapter 4, we classified existing methods by those that filled WM lesions using the *local* intensities from the surrounding neighboring voxels of lesions, and those that used *global* WM intensities from the whole brain to fill WM lesions. Although all these methods had been already validated separately, we performed a general comparison of all the available techniques in order to analyze their accuracy on the same 1.5T and 3T data and also to investigate its performance with different tissue segmentation techniques such as FAST and SPM8.

This analysis served as a basis to propose a new technique to refill WM lesions which was a compromise between *global* and *local* methods. In contrast to other existing techniques, the proposed method filled lesion voxels intensities with random values of a normal distribution generated from the mean WM signal intensity of each two-dimensional slice. Our results showed that when compared to other methods, our approach yielded the lowest deviation in GM and WM volume on 1.5T and 3T data when FAST tissue segmentation was used. When SPM8 tissue segmentation method was used, the performance of our lesion filling method was also very competitive, yielding the lowest differences or similar to the best method in GM and WM. In contrast to the rest of pipelines, differences in tissue volume between the same images filled with our method and afterwards segmented with either FAST or SPM8 were very low (< 0.1%), which indicates that the proposed strategy was equally efficient independently of the tissue segmentation chosen.

The proposed algorithm performed significantly better than local methods on images with higher lesion load. In contrast to *global* methods, *local* methods may be limited by the range of similar intensities coming from the neighboring voxels, which on images with a large lesions may be introducing a bias on GM and WM tissue distributions by the addition of a considerable number of voxels with similar intensities. Furthermore, the performance our approach was also better on images with high lesion load when compared with *local* methods, specially on images with lower resolution such as 1.5T data, most probably because our method estimated the mean global normal appearing WM intensity for each slice independently, being more sensible to reproduce possible changes in the intensities between slices.

7.3 Effect of automating lesion segmentation and filling on tissue segmentation

As already said earlier, lesion filling has been shown to be an effective method to reduce the effects of these lesions on tissue segmentation. However, in all the lesion-filling approaches including ours, MS lesions have to be known *a priori*, which requires to delineate lesions manually. This was a clear limitation in terms of fully automatizing brain tissue on the presence of MS lesions, which motivated the evaluation of the effect of automated lesion segmentation on tissue segmentation. Although different automated tissue segmentation methods have been proposed, most of them are based on supervised learning, which require to explicitly train them usually with a large amount of labeled data. Labeled data may be not available, which had pointed out the interest of the community in unsupervised methods that can operate without prior data. As shown in Chapter 5, we compared two fully unsupervised pipelines that combined automated lesion segmentation and filling as a first step to understand the effect of fully processed images in tissue segmentation.

Given the performance shown in our previous studies and its widely use in clinical studies, SPM8 was used as a reference tissue segmentation method to measure tissue volume on a set of 70 3T images of CIS patients. On these images, available manual expert annotations were employed to refill WM lesions before segmentation using the filling method of the pipeline, and were considered as ground-truth. Afterwards, we evaluated the differences in GM and WM volume between the set of filled images using manual annotations and the same images processed using different variations of the SLS and LST toolkits that differed in the level of manual intervention. Evaluating different pipelines with distinct levels of automation permitted us to analyze the effect of each of the automated process involved in the obtained differences in total and normal-appearing tissue volume.

As already suggested in Chapter 3, this new analysis showed that the effect of lesions in total tissue volume was limited due to a canceling effect between the errors produced in the same lesion regions, and the effect of these voxels in normal-appearing tissue. In all the pipelines that incorporated automated lesion segmentation, most of the observed differences in normal-appearing tissue were produced by the effect of false positive lesion voxels that were already segmented without refilling them. In contrast, there was not a relevant correlation between the number of false positive lesion voxels and the observed differences in normal-appearing GM and WM, which suggested that most of these miss-classified voxels were actually WM before refilling them. The relationship between errors in automated lesion and tissue segmentation suggest also the importance of not only to keep reducing the number of missed lesions, but also stress the necessity of contextual spatial information of lesion regions in order to confine them in WM and hence reduce the effect of

miss-classified voxels on tissue segmentation.

As shown in the results presented in Chapter 4, masking-out lesion voxels before tissue segmentation might not be optimal, as leaving lesion voxels out of the tissue distributions appears to increase the differences in tissue volume with respect to lesion filled images, even if these voxels are re-assigned to WM afterwards. However, although not optimal, masking lesion before segmentation has been found a valid alternative to reduce the effects of WM lesions in research and clinical settings, and only in the recent years, lesion filling techniques are been already applied on research and clinical studies. Regarding to this, our results show that at least with the evaluated data, the differences in tissue volume between images where expert lesion masks have been masked-out and the same images where lesions have been automatically segmented and filled, are similar on images with low lesion load ($< 10ml$). In contrast, from our experiments we observed that differences in tissue volume tend to increase with lesion load on masked-out images, while the increase of the error is more moderated on the fully-automated images. However, given the available data and maximum lesion load considered in our analysis ($< 20ml$), these findings should be considered with care.

In any case, our analysis points out the fact that automated lesion segmentation and filling methods reduced significantly the impact of WM lesions on tissue segmentation, and with a similar performance to the pipelines that required manual expert intervention. These results are relevant and validate that each of these automated processes can be useful not only in terms of time and economic costs, but also as active processes in fully automated tissue segmentation in the presence of WM lesions.

7.4 Fully automated tissue segmentation of images containing WM lesions

Previous sections have stressed the necessity to deal with MS lesions before tissue segmentation, showing several general insights that can be useful for automated tissue segmentation of images containing lesions. The obtained results of the different evaluated methods in Chapter 2 and 3 have pointed out the superiority of methods that were benefited by morphological prior information or spatial constraints in automated brain tissue segmentation. More importantly, the results obtained in Chapter 3 have evidenced the effect of WM lesion on tissue segmentation and the necessity to deal with MS lesions in order to reduce not only the bias produced by the same lesions but also the effect of these lesion voxels in normal-appearing tissue. In this scenario, we have proposed a new lesion filling technique that was very competitive with different databases and tissue segmentation methods, as shown in

Chapter 4. Finally, we have shown in Chapter 5 that the addition of unsupervised lesion segmentation and filling into already existing tissue segmentation pipelines reduced significantly the error in tissue volume when compared with previous pipelines where lesions were segmented as normal tissue.

Following these insights, we have developed a novel multi-channel method designed to segment brain tissues in MRI images of MS patients. As explained in Chapter 6, this approach makes use of a combination of intensity, anatomical and morphological prior maps to guide the tissue segmentation. Tissue segmentation has been tackled based on a robust partial volume segmentation where WM outliers have been estimated and refilled before segmentation using a multi-channel post-processing algorithm also integrating partial volume segmentation, spatial context, and prior anatomical and morphological atlases. Furthermore, the proposed method takes advantage of new affordable processors such as GPUs that reduce up to four times the execution time to register and segment tissue when compared to general purpose CPUs. This property makes this method useful for studies containing a large number of subjects to analyze.

The proposed method has been quantitatively and qualitatively evaluated using different databases of images containing WM lesions. In order to analyze the extent to which T1-w and FLAIR modalities intervened in the obtained accuracy, the proposed method was run in all experiment using only T1-w or using both T1-w and FLAIR image sequences. As shown by the presented results, the proposed technique yielded competitive and consistent results in both general and MS specific databases without parameter tweaking. In the MRBrainS tissue segmentation challenge¹, our method combining both T1-w and FLAIR was the best non-supervised technique of the challenge, being ranked in the 7th position out of 31 participant methods. When only the T1-w modality was used, still the accuracy of the proposed method was clearly superior to general purpose methods such as FAST (ranked 21th) and SPM12 (best ranked 17th), even if those used both image modalities. In MS data, the performance of our method combining T1-w and FLAIR sequences was similar or better to the best evaluated pipeline incorporating lesion segmentation and filling. Obtained differences in tissue volume between images processed with the proposed algorithm and the same images where lesions were filled using expert lesion annotations, were lower than 0.15% on all tissues, validating the overall capability of the proposed method to reduce the effects of WM lesions on tissue segmentation.

In general, our results showed that the percentages of error in tissue volume of our proposed approach were low and similar in both databases. The percentages of error were the lowest when the FLAIR modality was used, which evidences that this image sequence has a direct effect on the efficiency of the algorithm, and consequently it should be used when available. However, the accuracy of the method using only

¹<http://mrbrains13.isi.uu.nl/>

the T1-w modality was also superior to other general purpose strategies, which also evidences that the improvement in tissue segmentation was not only generated by the addition of the FLAIR modality, but also by the combination of intensity, anatomical and morphological priors, and the use of an specific outlier algorithm with integrated lesion filling.

Chapter 8

Conclusions

This thesis synthesizes our work done during the last three years. Following the same objectives defined in the Introduction chapter, we summarize in what follows the main conclusions and contributions of this thesis:

- We analyzed and evaluated the state-of-the-art of brain tissue segmentation methods. This first stage aimed to quantitative review and evaluate different proposed tissue segmentation techniques in order to understand their advantages and drawbacks. As part of the resulting analysis published in the *Journal of Magnetic Resonance Imaging* in January of 2015, **our results showed a higher accuracy on several methods that incorporated morphological prior information and/or spatial constraints such as FAST, SP-VASEG and SPM8. These methods were also less prone to changes in acquisition sequences and intensity inhomogeneities.**
- We studied the effect of WM lesions on tissue segmentation of MS patient images. The second stage to cover was focused on the analysis of the effects of WM lesions on the tissue distributions. Six of the analyzed methods on Chapter 2 were evaluated on multi-center 1.5T MS data from different scanners. Related to the previous stage, our results stressed **the necessity of adding morphological prior information and/or spatial constraints in automated brain tissue segmentation, not only to overcome inherent MRI artifacts but also as an important component to deal with WM lesions**. Furthermore, our analysis of the effects of WM lesions on tissue volume showed that **the inclusion of WM lesions on tissue segmentation not only biased the total tissue volume measurements by the addition of miss-classified lesion voxels, but also by the effect of these lesions in observed differences in normal-appearing tissue volume.**

The entire analysis was published in the *American Journal of Neuroradiology* in February 2015.

- We proposed a new technique to reduce the effects of WM lesions on tissue segmentation of MS patient images. The third stage required first to compare the accuracy of different proposed lesion filling techniques in the literature with the aim to propose then a new technique to reduce the effects of WM lesions on tissue segmentation. **The lesion filling method proposed in this thesis, was shown effective in different data and independently of the tissue segmentation method used afterwards.** The proposed approach outperformed the rest of methods on both 1.5T and 3T data when FAST was used, while its performance was similar or lower to the best available strategy when SPM8 was used. The proposed lesion filling method was published in the *NeuroImage: Clinical journal* in August of 2014. **Furthermore, we released a public version on the proposed method that can be freely downloaded from our research team web page¹.** This software is already been used in the collaborating hospitals.
- We analyzed and evaluated the effect of automated WM lesion segmentation and filling on the tissue segmentation. During the fourth stage proposed, we quantitatively evaluated the accuracy of two state-of-the-art automated pipelines that incorporate unsupervised lesion segmentation, lesion filling and tissue segmentation on MS data. As shown in the published paper in the *NeuroImage: Clinical journal* in October of 2015, our analysis showed that **pipelines that incorporated automated lesion segmentation and filling were capable to reduce significantly the impact of WM lesions on tissue segmentation, performing similarly to the pipelines that required manual expert intervention.**
- Finally, we proposed a new fully automated tissue segmentation method for MS patient images containing lesions. The main goal of the thesis was to propose a fully automated tissue segmentation method capable to deal with images of MS patients. As shown in Chapter 6, the proposed method incorporated all the major insights obtained from previous stages with the aim of provide a robust fully automated tissue approach for accurate brain volume measurements. Our results showed that **when compared with existent tissue segmentation methods, the presented approach yielded a higher accuracy in tissue segmentation while the influence of MS lesions on tissue segmentation was lower or similar to the best state-of-the-art pipeline incorporating automated lesion segmentation and filling.**

¹The latest version on the proposed lesion filling method can be download from <http://atc.udg.edu/nic/slftoolbox/index.html>

This work has been submitted for publication in the *Medical Image Analysis journal* in January 2016. **As part of this work, we also released a public version on the proposed method that can be freely downloaded from our research team web page².**

During this PhD thesis, different collaborations have been done with other researchers of the VICOROB group. First, we evaluated the effect of MS lesions on longitudinal registration in the published study of Diez et al. [23], where we contributed with several processing steps such as lesion filling. Lately, we were also involved in the development of several automated lesion segmentation pipelines that allowed us to gain knowledge about this topic. In this regard, we helped to implement two different lesion segmentation pipelines for MS such as those published in the papers of Cabezas et al. [10] and Roura et al. [53], respectively. Furthermore, we also collaborated on a new pipeline for automated lesion segmentation of Lupus lesions proposed by Roura et al., which has been submitted for publication recently.

8.1 Future work

Unfortunately, there are several aspects that have been not investigated during this thesis. One of the main limitations on several stages has been the lack of 3T images with high lesion load. As pointed out in Chapters 5 and 6, the low mean lesion load of the cohorts analyzed, which indeed has the major interest for the medical experts, has not allowed to investigate better the performance of the analyzed pipelines in the presence of images with higher lesion load. In the case of our tissue segmentation method, we believe that an additional analysis of the performance with images with higher lesion load would be helpful not only to analyze the robustness of the proposed algorithm, but also to investigate the benefits of adding other image sequences such as T2-w or PD-w.

Although the proposed tissue segmentation method has been designed for cross-sectional data, there is an increasingly clinical interest in the measurements of longitudinal changes in tissue volume. We believe that the proposed method could be extended to longitudinal changes by re-adapting the pipeline with prior registering of time point images before tissue segmentation. This is in fact one of the goals that our team has in mind to tackle first within the research framework of the BiomarkEM.cat project, in order to release suitable tools that can be used in clinical settings.

²A public version of the method can be download from <http://atc.udg.edu/nic/msseg/index.html>

The ultimate goal should be to provide state-of-the-art tools for the collaborating hospitals involved in these research projects that may be useful not only to diagnose and monitorize the progression of disease, but also to evaluate new treatments in MS patients. Related to that, the tools developed in this thesis should be integrated with other tools developed in our group in order to implement this complete system capable to provide robust useful biomarkers in MS such as the number of lesions, lesion volume, brain tissue volume or brain atrophy.

Bibliography

- [1] Julio Acosta-Cabronero, Guy B. Williams, João M S Pereira, George Pengas, and Peter J. Nestor. The impact of skull-stripping and radio-frequency bias correction on grey-matter segmentation for voxel-based morphometry. *NeuroImage*, 39(4):1654–1665, 2008.
- [2] Ayelet Akselrod-Ballin, Meirav Galun, Moshe John Gomori, Ronen Basri, and Achi Brandt. Atlas guided identification of brain structures by combining 3D segmentation and SVM classification. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 9(Pt 2):209–16, 2006.
- [3] J B Arnold, J S Liow, K a Schaper, J J Stern, J G Sled, D W Shattuck, a J Worth, M S Cohen, R M Leahy, J C Mazziotta, and D a Rottenberg. Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects. *NeuroImage*, 13(5):931–943, 2001.
- [4] John Ashburner and Karl J. Friston. Unified segmentation. *NeuroImage*, 26:839–851, 2005.
- [5] Marco Battaglini, Mark Jenkinson, Nicola De Stefano, and Nicola De Stefano. Evaluating and reducing the impact of white matter lesions on brain volume measurements. *Human Brain Mapping*, 33(9):2062–71, 2012.
- [6] S. Bricq, Ch. Collet, and J.P. Armspach. Unifying framework for multimodal brain mri segmentation based on hidden markov chains. *Medical Image Analysis*, 12(6):639 – 652, 2008. *|ce:title|Special issue on information processing in medical imaging 2007|ce:title|*.
- [7] Per Brodal. *The central nervous system*. 2010.
- [8] Mariano Cabezas, Arnau Oliver, Xavier Lladó, Jordi Freixenet, and Meritxell Bach Cuadra. A review of atlas-based segmentation for magnetic resonance brain images. *Computer Methods and Programs in Biomedicine*, 104(3):e158–e177, 2011.

- [9] Mariano Cabezas, Arnau Oliver, Eloy Roura, Jordi Freixenet, Joan C. Vilanova, Lluís Rami-Torrent, ??lex Rovira, and Xavier Lladó. Automatic multiple sclerosis lesion detection in brain MRI by FLAIR thresholding. *Computer Methods and Programs in Biomedicine*, 115(3):147–161, 2014.
- [10] Mariano Cabezas, Arnau Oliver, Sergi Valverde, Brigitte Beltran, Jordi Freixenet, Joan C. Vilanova, Lluís Rami-Torrent, ??lex Rovira, and Xavier Lladó. BOOST: A supervised approach for multiple sclerosis lesion segmentation. *Journal of Neuroscience Methods*, 237:108–117, 2014.
- [11] Benoît Caldairou, Nicolas Passat, Piotr A Habas, Colin Studholme, and François Rousseau. A non-local fuzzy segmentation method: application to brain mri. *Pattern Recognition*, 44(9):1916–1927, 2011.
- [12] A Ceccarelli, J S Jackson, S Tauhid, A Arora, J Gorky, E Dell’Oglio, A Bakshi, T Chitnis, S J Khouri, H L Weiner, C R G Guttmann, R Bakshi, and M Neema. The impact of lesion in-painting and registration methods on voxel-based morphometry in detecting regional cerebral gray matter atrophy in multiple sclerosis. *American Journal of Neuroradiology*, 33(8):1579–85, September 2012.
- [13] Antonia Ceccarelli, Maria a Rocca, Mohit Neema, Vittorio Martinelli, Ashish Arora, Shahamat Tauhid, Angelo Ghezzi, Giancarlo Comi, Rohit Bakshi, and Massimo Filippi. Deep gray matter T2 hypointensity is present in patients with clinically isolated syndromes suggestive of multiple sclerosis. *Multiple sclerosis (Hounds Mills, Basingstoke, England)*, 16(1):39–44, 2010.
- [14] D T Chard, G J M Parker, R Kapoor, A J Thompson, and D H Miller. Brain atrophy in clinically early relapsing-remitting multiple sclerosis. 2002.
- [15] Declan T. Chard, Jonathan S. Jackson, David H. Miller, and Claudia A M Wheeler-Kingshott. Reducing the impact of white matter lesions on automated measures of brain gray and white matter volumes. *Journal of Magnetic Resonance Imaging*, 32:223–228, 2010.
- [16] Harvey E. Cline, William E. Lorenzen, Ron Kikinis, and Ferenc Jolesz. Three-Dimensional Segmentation of MR Images of the Head Using Probability and Connectivity. *Journal of Computer Assisted Tomography*, 14(6):1037–1045, 1990.
- [17] Alastair Compston and Alasdair Coles. Multiple sclerosis. *Lancet*, 359(9313):1221–31, 2002.
- [18] Alastair Compston and Alasdair Coles. Multiple sclerosis. *Lancet*, 372(9648):1502–17, 2008.

- [19] Devic M Confavreux C, Aimard G. Course and prognosis of multiple sclerosis assessed by the computerized data processing of 349 patients. *Brain*, 103(2):281–300, 1980.
- [20] Renske De Boer, Henri A Vrooman, Fedde Van Der Lijn, Meike W Vernooij, M Arfan Ikram, Aad Van Der Lugt, Monique M B Breteler, and Wiro Niessen. White matter lesion extension to automatic brain tissue segmentation on mri. *NeuroImage*, 45(4):1151–1161, 2009.
- [21] N. De Stefano, A. Giorgio, M. Battaglini, M. Rovaris, M. P. Sormani, F. Barkhof, T. Korteweg, C. Enzinger, F. Fazekas, M. Calabrese, D. Dinacci, G. Tedeschi, A. Gass, X. Montalban, A. Rovira, A. Thompson, G. Comi, D. H. Miller, and M. Filippi. Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes. *Neurology*, 74(23):1868–1876, 2010.
- [22] Hrishikesh Deshpande, Pierre Maurel, and Christian Barillot. Classification of Multiple Sclerosis Lesions using Adaptive Dictionary Learning. *Computerized Medical Imaging and Graphics*, 2015.
- [23] Yago Diez, Arnau Oliver, Mariano Cabezas, Sergi Valverde, Robert Martí, Joan Carles Vilanova, Lluís Ramió-Torrentà, Àlex Rovira, and Xavier Lladó. Intensity based methods for brain MRI longitudinal registration. A study on multiple sclerosis patients. *Neuroinformatics*, 12(3):365–379, 2014.
- [24] Robert R Edelman and Steven Warach. Magnetic Resonance Imaging. *New England Journal of Medicine*, 328(10):708–716, 1993.
- [25] Massimo Filippi, Paolo Preziosa, Massimiliano Copetti, Gianna Riccitelli, Mark a Horsfield, Vittorio Martinelli, Giancarlo Comi, and Maria a Rocca. Gray matter damage predicts the accumulation of disability 13 years later in MS. *Neurology*, 81(20):1759–67, nov 2013.
- [26] Massimo Filippi and Maria A Rocca. MR imaging of multiple sclerosis. *Radiology*, 259(3):659–81, 2011.
- [27] Elizabeth Fisher, Jar-Chi Lee, Kunio Nakamura, and Richard A Rudick. Gray matter atrophy in multiple sclerosis: A longitudinal study. *Annals of Neurology*, 64(3):255–265, sep 2008.
- [28] Leonora K Fisniku, Declan T Chard, Jonathan S Jackson, Valerie M Anderson, Daniel R Altmann, Katherine A Miszkiel, Alan J Thompson, and David H Miller. Gray Matter Atrophy Is Related to Long- Term Disability in Multiple Sclerosis. *Ann Neurol*, 64:247–254, 2008.

- [29] Onur Ganiler, Arnau Oliver, Yago Diez, Jordi Freixenet, Joan C. Vilanova, Brigitte Beltran, Llu??s Rami??-Torrent??, ??lex Rovira, and Xavier Llad?? A subtraction pipeline for automatic detection of new appearing multiple sclerosis lesions in longitudinal studies. *Neuroradiology*, 56(5):363–374, 2014.
- [30] Daniel Garc??a-Lorenzo, Simon Francis, Sridar Narayanan, Douglas L Arnold, and D Louis Collins. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Medical Image Analysis*, 17(1):1–18, January 2013.
- [31] Ezequiel Geremia, Olivier Clatz, Bjoern H. Menze, Ender Konukoglu, Antonio Criminisi, and Nicholas Ayache. Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images. *NeuroImage*, 57(2):378–390, 2011.
- [32] Guido Gerig, John Martin, Ron Kikinis, Olaf Kubler, Martha Shenton, and Ferenc A. Jolesz. Unsupervised tissue type segmentation of 3D dual-echo MR head data. *Image and Vision Computing*, 10(6):349–360, 1992.
- [33] Tal Geva. Magnetic resonance imaging: historical perspective. *Journal of Cardiovascular Magnetic Resonance*, 8(4):573–580, 2006.
- [34] Antonio Giorgio and Nicola De Stefano. Clinical use of brain volumetry. *Journal of Magnetic Resonance Imaging*, 37(1):1–14, 2013.
- [35] Nicolas Guizard, Pierrick Coup??, Vladimir S. Fonov, Jose V. Manj??n, Douglas L. Arnold, and D. Louis Collins. Rotation-invariant multi-contrast non-local means for MS lesion segmentation. *NeuroImage: Clinical*, 8:376–389, 2015.
- [36] Rola Harmouche, Nagesh K Subbanna, D Louis Collins, Douglas L Arnold, and Tal Arbel. Based on Modeling Regional Intensity Variability and Local Neighborhood Information. 62(5):1281–1292, 2015.
- [37] Zujun Hou. A review on MR image intensity inhomogeneity correction, 2006.
- [38] Saurabh Jain, Diana M. Sima, Annemie Ribbens, Melissa Cambron, Anke Maertens, Wim Van Hecke, Johan De Mey, Frederik Barkhof, Martijn D. Steenwijk, Marita Daams, Frederik Maes, Sabine Van Huffel, Hugo Vrenken, and Dirk Smeets. Automatic segmentation and volumetry of multiple sclerosis brain lesions from MR images. *NeuroImage: Clinical*, 8:367–375, 2015.
- [39] T Kapur, W E Grimson, W M Wells, and R Kikinis. Segmentation of brain tissue from magnetic resonance images. *Medical image analysis*, 1(2):109–27, 1996.

- [40] Jong Min Lee, Uicheul Yoon, Sang Hee Nam, Jung Hyun Kim, In Young Kim, and Sun I. Kim. Evaluation of automated and semi-automated skull-stripping algorithms using similarity index and segmentation error. *Computers in Biology and Medicine*, 33(6):495–507, 2003.
- [41] Xavier Lladó, Onur Ganiler, Arnau Oliver, Robert Martí, Jordi Freixenet, Laia Valls, Joan C. Vilanova, Lluís Ramió-Torrentà, and Àlex Rovira. Automated detection of multiple sclerosis lesions in serial brain MRI. *Neuroradiology*, 54:787–807, 2012.
- [42] Xavier Lladó, Arnau Oliver, Mariano Cabezas, Jordi Freixenet, Joan C. Vilanova, Ana Quiles, Laia Valls, Lluís Ramió-Torrentà, and Àlex Rovira. Segmentation of multiple sclerosis lesions in brain MRI: A review of automated approaches. *Information Sciences*, 186(1):164–185, March 2012.
- [43] F. D. Lublin and S. C. Reingold. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. *Neurology*, 46(4):907–911, April 1996.
- [44] Stefano Magon, Laura Gaetano, M Mallar Chakravarty, Jason P Lerch, Yvonne Naegelin, Christoph Stippich, Ludwig Kappos, Ernst-Wilhelm Radue, and Till Sprenger. White matter lesion filling improves the accuracy of cortical thickness measurements in multiple sclerosis patients: a longitudinal study. *BMC Neuroscience*, 15(1):106, January 2014.
- [45] D Mahapatra. Analyzing Training Information From Random Forests for Improved Image Segmentation. *Image Processing, IEEE Transactions on*, 23(4):1504–1512, 2014.
- [46] J.L. Marroquin, B.C. Vemuri, S. Botello, E. Calderon, and A. Fernandez-Bouzas. An accurate and efficient bayesian method for automatic segmentation of brain mri. *Medical Imaging, IEEE Transactions on*, 21(8):934 –945, aug. 2002.
- [47] Dzung L Pham. Spatial Models for Fuzzy Clustering. *Computer Vision and Image Understanding*, 297:285–297, 2001.
- [48] Chris H Polman, Stephen C Reingold, Brenda Banwell, Michel Clanet, Jeffrey a Cohen, Massimo Filippi, Kazuo Fujihara, Eva Havrdova, Michael Hutchinson, Ludwig Kappos, Fred D Lublin, Xavier Montalban, Paul O’Connor, Magnhild Sandberg-Wollheim, Alan J Thompson, Emmanuelle Waubant, Brian Weinshenker, and Jerry S Wolinsky. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of neurology*, 69(2):292–302, 2011.

- [49] V. Popescu, M. Battaglini, W. S. Hoogstrate, S. C J Verfaillie, I. C. Sluimer, R. A. van Schijndel, B. W. van Dijk, K. S. Cover, D. L. Knol, M. Jenkins, F. Barkhof, N. de Stefano, H. Vrenken, F. Barkhof, X. Montalban, F. Fazekas, M. Filippi, J. Frederiksen, L. Kappos, D. Miller, J. Palace, C. Polman, M. Rocca, A. Rovira, and T. Yousry. Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis. *NeuroImage*, 61(4):1484–1494, 2012.
- [50] V. Popescu, N. C G Ran, F. Barkhof, D. T. Chard, C. A. Wheeler-Kingshott, and H. Vrenken. Accurate GM atrophy quantification in MS using lesion-filling with co-registered 2D lesion masks. *NeuroImage: Clinical*, 4:366–373, 2014.
- [51] Martin Rajchl, John S.H. Baxter, a. Jonathan McLeod, Jing Yuan, Wu Qiu, Terry M. Peters, and Ali R. Khan. Hierarchical max-flow segmentation framework for multi-atlas segmentation with Kohonen self-organizing map based Gaussian mixture modeling. *Medical Image Analysis*, 000:1–12, 2015.
- [52] E. Roura, T. Schneider, M. Modat, P. Daga, N. Muhlert, D. Chard, S. Ourselin, X. Lladó, and C. A. M. Wheeler-Kingshott. Multi-channel registration of fa and t1w images in the presence of atrophy: application to multiple sclerosis. *Functional Neurology*, 30(4), 2015.
- [53] Eloy Roura, Arnau Oliver, Mariano Cabezas, Sergi Valverde, Deborah Pareto, Joan C Vilanova, Lluís Ramió-Torrentà, Àlex Rovira, and Xavier Lladó. A toolbox for multiple sclerosis lesion segmentation. *Neuroradiology*, 57(10):1031–1043, 2015.
- [54] Eloy Roura, Arnau Oliver, Mariano Cabezas, Joan C. Vilanova, ??lex Rovira, Lluís Ramió-Torrentà, and Xavier Lladó. MARGA: Multispectral Adaptive Region Growing Algorithm for brain extraction on axial MRI. *Computer Methods and Programs in Biomedicine*, 113(2):655–673, 2014.
- [55] Snehashis Roy, Aaron Carass, Pierre-Louis Bazin, Susan Resnick, and Jerry L. Prince. Consistent segmentation using a rician classifier. *Medical Image Analysis*, 16(2):524 – 535, 2012.
- [56] Richard a Rudick, Jar-Chi Lee, Kunio Nakamura, and Elizabeth Fisher. Gray matter atrophy correlates with MS disability progression measured with MSFC but not EDSS. *Journal of the neurological sciences*, 282(1-2):106–11, jul 2009.
- [57] Paul Schmidt, Christian Gaser, Milan Arsic, Dorothea Buck, Annette Förtschler, Achim Berthele, Muna Hoshi, Rüdiger Ilg, Volker J Schmid, Claus Zimmer, Bernhard Hemmer, and Mark Mühlau. An automated tool for detection of

- FLAIR-hyperintense white-matter lesions in Multiple Sclerosis. *NeuroImage*, 59(4):3774–3783, February 2012.
- [58] Michaël Sdika and Daniel Pelletier. Nonrigid registration of multiple sclerosis brain images using lesion inpainting for morphometry or lesion mapping. *Human Brain Mapping*, 30(4):1060–1067, 2009.
 - [59] D W Shattuck, S R Sandor-Leahy, K a Schaper, D a Rottenberg, and R M Leahy. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13(5):856–876, 2001.
 - [60] A. Simmons, P. S. Tofts, G. J. Barker, and S. R. Arridge. Sources of intensity nonuniformity in spin echo images at 1.5 T. *Magnetic Resonance in Medicine*, 32(1):121–128, 1994.
 - [61] J G Sled, A P Zijdenbos, and A C Evans. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE transactions on medical imaging*, 17(1):87–97, 1998.
 - [62] Stephen M. Smith. Fast robust automated brain extraction. *Human Brain Mapping*, 17(3):143–155, 2002.
 - [63] G H Sperber. Clinically Oriented Anatomy. *Journal of anatomy*, 208(3):393, 2006.
 - [64] Lawrence Steinman, M.D. Multiple Sclerosis: A Coordinated Immunological Attack against Myelin in the Central Nervous System. *Cell*, 85(3):299–302, may 1996.
 - [65] Elizabeth M Sweeney, Russell T Shinohara, Navid Shiee, Farrah J Mateen, Avni a Chudgar, Jennifer L Cuzzocreo, Peter a Calabresi, Dzung L Pham, Daniel S Reich, and Ciprian M Crainiceanu. OASIS is Automated Statistical Inference for Segmentation, with applications to multiple sclerosis lesion segmentation in MRI. *NeuroImage: Clinical*, 2:402–13, jan 2013.
 - [66] Jussi Tohka, Ivo D. Dinov, David W. Shattuck, and Arthur W. Toga. Brain mri tissue classification based on local markov random fields. *Magnetic Resonance Imaging*, 28(4):557 – 573, 2010.
 - [67] Xavier Tomas-Fernandez and Simon Keith Warfield. A Model of Population and Subject (MOPS) Intensities with Application to Multiple Sclerosis Lesion Segmentation. *IEEE transactions on medical imaging*, 0062(c):1–15, jan 2015.
 - [68] Nicholas J. Tustison, Brian B. Avants, Philip A. Cook, Yuanjie Zheng, Alexander Egan, Paul A. Yushkevich, and James C. Gee. N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6):1310–1320, 2010.

- [69] Annegreet van Opbroek, Fedde van der Lijn, and Marleen de Bruijne. Automated brain-tissue segmentation by multi-feature svm classification. In *Proceedings of the MICCAI Workshopsâ€The MICCAI Grand Challenge on MR Brain Image Segmentation (MRBrainSâ€13)*, 2013.
- [70] H Vrooman, Fedde van der Lijn, and W Niessen. Auto-knn: brain tissue segmentation using automatically trained knearest-neighbor classification. In *Proceedings of the MICCAI Workshopsâ€The MICCAI Grand Challenge on MR Brain Image Segmentation (MRBrainSâ€13)*, 2013.
- [71] Zhao Yi, Antonio Criminisi, Jamie Shotton, and Andrew Blake. Discriminative, semantic segmentation of brain tissue in mr images. In Guang-Zhong Yang, David Hawkes, Daniel Rueckert, Alison Noble, and Chris Taylor, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009*, volume 5762 of *Lecture Notes in Computer Science*, pages 558–565. Springer Berlin / Heidelberg, 2009.
- [72] Y Zhang, M Brady, and S Smith. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20:45–57, 2001.