Partial Volume Segmentation in Magnetic Resonance Imaging (MRI)

By:

Jonás Grande Barreto

A dissertation submitted in partial fulfillment
of the requirements for the degree of:

Ph.D. IN COMPUTER SCIENCE

at

Instituto Nacional de Astrofísica, Óptica y Electrónica

December, 2020
Tonantzintla, Puebla, México

Supervised by:

Ph.D. María del Pilar Gómez Gil, INAOE

Abstract

Brain tissue segmentation using magnetic resonance imaging (MRI) rises as an essential tool for the human brain's clinical analysis. The advances in brain MRI provide a large amount of data with high spatial resolution. However, different imaging device settings produce different imaging conditions, causing artifacts such as noise and inhomogeneity. Irregular shapes of the tissues and the Partial volume effect are challenges to overcome during the brain tissues' segmenting. Hence, brain tissue segmentation becomes a tedious (time-consuming) task for clinicians who manually select relevant information and frequently prone to errors. We propose a new hybrid computational model based on fuzzy strategies and artificial neural networks for automatic brain tissue segmentation in this work. The model divides into two phases. The first phase fits a generic atlas to a subject target. It is done by combining 3D feature descriptors and atlas information applying a novel fuzzy function. The proposed method, named Gardens2, estimates each voxel's membership degree of the fitted atlas model. The second phase carries out the brain tissue segmentation. We use pseudo-labels, computed from the output of Gardens2, to support inter-class separation and an intra-class compactness. Pseudo-labels are calculated using an extension of the Self-Organized Map (SOM) training algorithm, named *Pseudo-Label Assisted Self-Organized Map* (PLA-SOM), which incorporates a novel mapping strategy. This novel strategy uses the summation of the memberships from the neighbors of the best-matching unit (topological information), associated with the input pattern, is used to define the pattern's label. This new mapping approach was practical for dealing with uncertainty in border prototypes. The proposed model was validated using F1-measure on three neuroimaging datasets: BrainWeb, IBSR18, and IBSR20. These datasets contain brain MRI volumes from healthy caucasian subjects. The experimental results revealed that, in most cases, our model outperformed the most popular

brain tissue segmentation models.

Resumen

La segmentación del tejido cerebral mediante imágenes por resonancia magnética surge como una herramienta esencial para el análisis clínico del cerebro humano. Los avances en los estudios por resonancia magnética cerebral proporcionan una gran cantidad de datos de alta resolución espacial. Sin embargo, diferentes configuraciones en los dispositivos de captura producen diferentes condiciones de salida en la imagen final, lo que hace que aparezcan artefactos como ruido y falta de homogeneidad. Las formas irregulares de los tejidos cerebrales y el efecto de volumen parcial son desafíos a superar durante la segmentación de los tejidos cerebrales. Por lo tanto, la segmentación del tejido cerebral se convierte en una tarea tediosa (que demanda mucho tiempo) para los especialistas clínicos que seleccionan manualmente la información relevante, y con frecuencia son propensos a errores.

En este trabajo proponemos un nuevo modelo computacional híbrido para la segmentación automática del tejido cerebral, basado en estrategias difusas y redes neuronales artificiales y que consta de dos fases. La primera fase ajusta un atlas genérico a una imagen objetivo. Esta tarea se realiza mediante la combinación de descriptores de características 3D e información extraída del atlas, aplicando una función difusa novedosa. El método propuesto, llamado Gardens2, estima el grado de pertenencia de cada vóxel del modelo del atlas ajustado. La segunda fase lleva a cabo la segmentación del tejido cerebral. En esta etapa, se utilizan pseudo-etiquetas, calculadas a partir de la salida de Gardens2, para ayudar a la separación entre clases y a la compacidad de cada clase. Las pseudo-etiquetas se calculan usando una extensión del algoritmo *Self-Organized Map* (SOM), llamado *Pseudo-Label Assisted Self-Organized Map* (PLA-SOM), el cual incorpora una nueva estrategia de mapeo. Esta novedosa estrategia utiliza la suma de las membresías de los vecinos alrededor de la mejor unidad de coincidencia (información topológica), asociada con el patrón

de entrada, para definir la etiqueta del patrón. Este nuevo enfoque de mapeo resultó práctico para lidiar con la incertidumbre en los prototipos de fronteras.

El modelo propuesto se validó utilizando *F1-measure* en tres conjuntos de datos de neuroimagen: BrainWeb, IBSR18 e IBSR20. Estos conjuntos contienen volúmenes de resonancia magnética cerebral de sujetos caucásicos sanos. Los resultados experimentales revelaron que, en la mayoría de los casos, nuestro modelo superó a los modelos de segmentación de tejido cerebral más populares.

Palabras clave: segmentación cerebral, MRI, atlas cerebral, membresías difusas, SOM, PLA-SOM.

LIST OF ACRONYMS

BET   Brain Extraction Tool

BMU   Best Matching Unit

BSE   Brain Surface Extractor

CMA   Center for Morphometric Analysis

CSF   Cerebrospinal Fluid

DSC   Dice Similarity Coefficient

EM   Expectation Maximization

EGS-SOM   Entropy-Gradient Segmentation

FAST   FMRIB's Automated Segmentation Tool

FCM   Fuzzy C-Means

FANTASM   Fuzzy and Noise Tolerant Adaptive Segmentation Method

FWMH   Full width at half maximum

GMM   Gaussian Mixture Model

GM   Gray Matter

HFS-SOM   Histogram based Fast Segmentation

IBSR   Internet Brain Segmentation Repository

ICBM   International Consortium of Brain Mapping

INU   Intensity Inhomogeneity

k-NN	k-Nearest-Neighbor
LVQ	Learning Vector Quantization
LGMM	Local Gaussian Mixture Models
MLP	Multi-Layer Perceptron
MRF	Markov Random Field
MAP	Maximum-A-Posteriori
MNI	Montreal Neurological Institute
MRI	Magnetic Resonance Imaging
NLRFCM	Non-Local Regularized FCM
PLA-SOM	Pseudo-Labeld Assisted SOM
PVC	Partial Volume Classifier
PVE	Partial Volume-Effect
PVS	Partial Volume Segmentation
ROI	Region Of Interest
RMSE	Root-Mean-Square Error
SOAC	Self-Organizing Active Contour
SOM	Self-Organized Map
TFE	Tissue Fraction Effect
US	Unified Segmentation
WM	White Matter

Notations

Ω	domain of the image
I_n	normalized image
I_o	raw image
$\omega \times \omega \times \omega$	sampling cube for co-occurrence matrix calculation
$S = \{S_1, \ldots, S_R\}$	watershed partition
S_i	subregion
$|\cdot|$	cardinality of the set
$\mathbf{X}$	universe of discourse
$\mathbf{x}_j$	j-th input pattern
x_{ijk_d}	denotes d-th feature of the k-th voxel in the j-th subregion of the i-th cluster
γ	a scalar parameter that regulates the amount of prior tissue knowledge
β	a scalar parameter that regulates amount of topological information from the SOM
$\Gamma = \{\Gamma_{csf}, \Gamma_{gm}, \Gamma_{wm}\}$	prior tissue model for CSF, GM, and WM
$\mathbf{w}_j$	prototype
$\widehat{y}$	pseudo label
$\mu_A(x)$	membership function
$P = \{P_{csf}, P_{gm}, P_{wm}\}$	membership level of each voxel to every class
η	learning rate parameter for SOM training
$\Lambda_{jk}(t)$	neighborhood function centered around the BMU$_j$
$\Delta \mathbf{w}_{ij}(t)$	weight correction at iteration t
B_k	neighbor prototypes centered on the BMU for the k-th voxel
$C = \{csf, gm, wm\}$	brain tissue clases
c_k	tissue class

Contents

List of Figures

LIST OF TABLES

Introduction

Magnetic resonance imaging (MRI) is a non-invasive medical imaging technique that produces images from inside of the human body. Its working principle involves subjecting a patient to an electromagnetic field to align hydrogen atoms in the same direction. Subsequently, a radio-frequency wave is induced, which disturbs the atoms in such a way that they start to resonate Buxton (2009). The relaxation time from such particles' resonance is measured by sensors sending signals to a computer, which converts them into a volume of images called *scans*. MRI plays an important role in assessing several health treatments, surgical planning, and medical research. Brain tissue segmentation in MRI is an important and challenging task in the analysis of the human brain.

Accurate brain tissue segmentation in MRI scans provides useful information for *in-vivo* quantization of tissues and their alterations due to brain growth, diagnosis of brain disorders, and surgical interventions Liu et al. (2019). Currently, manual segmentation is considered as the gold standard, but it is a time-consuming task. Besides, the associated rater bias compromises segmentation results Prakash and Kumari (2017); Makropoulos et al. (2018). Therefore, there is a need for accurate and automatic methods that provide reproducible results for segmenting the brain into regions or tissues of interest.

1.1 Motivation and justification

There is currently a wide range of automatic methods that have addressed the brain tissue segmentation problem Balafar et al. (2010); Despotović et al. (2015); Makropoulos et al. (2018). Nevertheless, an essential issue for those methods is related to the ambiguity defining the brain tissues' borders. This

ambiguity is generated by a phenomenon called *Partial Volume-Effect* (PVE), which makes that the gray value of a voxel in a brain MRI scan represents a mixture of several tissue types. In this research, we adopt the term *mixel*, proposed by Choi Choi et al. (1991), to refer to those voxels which contain multiple tissue classes Choi et al. (1991). Even using high-resolution devices, PVE is unavoidable because the problem moves to other zones on the image. Consequently, the PVE leads the models to sub-optimal segmentation results. Image segmentation techniques must estimate each tissue's contribution in a mixel to define the brain tissues' borders accurately and overcome this manner, the challenges produced by PVE.

After analyzing related work, we observed that most of the existing approaches search for incorporating *prior information* from a *brain atlas*, which provides spatial information and prior probability for each voxel to every tissue class. The prior information helps to bound the PVE and improves the segmentation performances of the methods. However, incorporating this kind of information is neither optimal in current methods because they are based on the analysis of edges and gray levels in the scan, which can be seriously affected by the bias field and PVE. These approaches force current methods to depend on the resolution of the scan and filtering processes, which help improve the contrast in the scan but generate a loss of information at the border of the objects.

1.2 Problem statement

In the brain tissue segmentation domain, PVE affects the structural representation of tissues and their characterization. As a result of this phenom, segmentation methods face uncertainty defining the boundaries between brain tissues because there is a set of voxels x that seems to belong to more than one tissue class c_k. This condition is a problem when a hard parcellation of brain tissues is required, and it increases its challenge when there is no labeled data to perform supervised learning. Thus, the brain tissue segmentation problem states as:

$$\{\exists x \mid x \in c_i \wedge x \in c_j, \quad i,j = 1,\ldots,C \text{ and } i \neq j\} \tag{1.1}$$

We address the outlined problem by proposing a computational model that takes advantage of available prior knowledge in a brain atlas and uses that to create pseudo-labels so that the inter-class separation and intra-class compactness of classes are promoted. A novel neuro-fuzzy approach is proposed to estimate each voxel's membership degree to every class and define the proper labels.

1.3 Contributions

This research focuses on the establishment of a solution to the ill-posed problem of segmentation in MRI such that the segmentation parcellates the brain MRI scan into three main tissues, named White Matter (WM), Gray Matter (GM), and Cerebrospinal Fluid (CSF).

The main contributions of this work are:

1. A procedure that fits a generic brain atlas to a specific target. We achieved this goal through the proposed Gardens2 algorithm.

2. An extension of the SOM algorithm, named PLA-SOM, where the topological information of the map and prior knowledge participate in the mapping phase.

3. A robust brain tissue segmentation pipeline.

1.4 Research question and Hypothesis

The research question that guide this research is:

- *How to accurately adjust a model, which has been estimated by an atlas registration process, to an specific target?*

Hypothesis:

An unsupervised computational model that bounds the impact of PVE by incorporating prior information and fuzzy memberships supports the development of a framework that obtains a better segmentation performance on brain MRI scans than whose models are based only on plain clustering.

The success condition will be measured in terms of the Dice Similarity Coefficient, also known as the F_1 measure Taha and Hanbury (2015).

1.5   Objectives

To develop and validate a novel computational model to overcome the challenges introduced by PVE on brain MRIs for achieving a precise segmentation of the brain tissues concerning to F_1 metric.

The specific objectives are:

- To develop a method to separate pure voxels from mixels on brain MRI scans.

- To develop a method based on fuzzy and neural clustering for classifying mixels into one of three possible main classes (CSF, GM, and WM) in brain tissues.

1.6   Rationale of the solution

The rationale behind proposing our method is that training a model to recognize mixed voxels in an image results in a complex task because, strictly speaking, all the elements have a mixing percentage. An alternative is first to identify the voxels with a stronger membership level towards the existing classes and train the model with those voxels' information. Subsequently, the trained model will be used to classify those voxels that have a higher mixing level. Therefore, our method divides the classification task into two phases: atlas adjustment and clustering. The atlas adjustment phase, executed by our proposed Gardens2 algorithm, fits a generic CSF, GM, and WM atlas label image (prior tissue models) to a target subject, using a novel fuzzy function combines spatial information and 3D texture features. Gardens2 algorithm is applied scan by scan to produce a precise adjustment between the prior tissue models and the target subject. We assumed that a better estimation of prior tissue maps drives to a better segmentation performance. We assessed the performance of Gardens2 against four well-established methods, namely BRAINS Johnson et al. (2007), FAST Zhang et al. (2001), US Ashburner and Friston (2005), and PVC Shattuck et al. (2001), and the results were promising.

The clustering phase, executed by our proposed PLA-SOM algorithm. Instead of cluster each input pattern directly to its BMU, we propose using

the membership summation from each prototype to every class to define the input pattern label. We assumed that including the data distribution coded in the prototypes and their topological sorting on the map could improve the mapping phase. The reason for doing this is because the prototypes close to the BMU learn to detect/reject the input patterns of a particular class. On the other hand, a single BMU can be related to more than one class, increasing the uncertainty of clustering new input patterns. Therefore, the uncertainty of mapping a new pattern can be reduced if the BMU's neighborhood also participates in the clustering o new input patterns.

1.7 Scope

Our method's scope is bounded to work with brain MRI scans from healthy average adults only, and it does not involve detecting pathologies in the brain. However, the proposed method can be applied as an in-direct aid in assessing pathologies based on brain volume reduction, such as in the progressive brain atrophy, which produces a decrease of the brain volume. For the atlas information, we consider probabilistic atlases constructed from the data of caucasian population since these are the most frequently data published in the literature.

1.8 Publications

At the point of the preparation of this manuscript, the following publications related to this work have been generated:

- Partial volume segmentation in magnetic resonance imaging (MRI). Tech. Rep. CCC-17-07.

- Unsupervised brain tissue segmentation in brain MRI images. In 2018 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC) Grande-Barreto and Gómez-Gil (2018).

- Segmentation of MRI brain scans using spatial constraints and 3D features. In Medical & Biological Engineering & Computing (MBEC) Grande-Barreto and Gómez-Gil (2020).

1.9   Organization of the document

The document is divided into six chapters. The following chapters corresponds to

Chapter 2 Background: This chapter presents the theoretical bases and the fundamental concepts for the development of the proposed research.

Chapter 3 State of the art: In this chapter, the related works on brain tissue segmentation are analyzed, and the contributions of our research are highlighted.

Chapter 4 The proposed method: This chapter describes the development of the proposed method, divided into two phases: fitting and segmentation. The fitting phase consists of adjusting a generic brain atlas to a specific target through the Gardens2 algorithm. The segmentation phase is performed with the proposed PLA-SOM method.

Chapter 5 Experimental results: This chapter presents the datasets and the metrics used to validate the proposed method's segmentation results. Additionally, the performance of Gardens2-PLA-SOM was compared against the performance of other methods specialized on brain tissue segmentation.

Chapter 6 Conclusions: This chapter summarizes the main findings of this doctoral research and describes future approaches to follow.

BACKGROUND

This chapter briefly introduces the theoretical concepts used in this research. The chapter divides into two parts; the first part states the definitions of brain tissue segmentation, Partial Volume Effect (PVE), and Partial Volume Segmentation; the second part introduces definitions and methods related to machine learning and image processing.

2.1 MRI fundamentals

Matter is made of atoms, where their nucleus is composed of protons and neutrons, and surrounded by an electron cloud. At the equilibrium state, the number of electrons is equal to the number of protons. MRI is based on electromagnetic effects in the nucleus of the atom. In the case of the human body, which is composed of about 70% of water, MRI measures hydrogen atoms' behavior; this is because a hydrogen atom consists of just one proton. These particles are in constant movement, for the hydrogen atom, its movement is a self-rotation (spin), which has a magnetic moment.

The above description must be translated from particles to tissues where the spins sum up to a macroscopic magnetic moment M. Suppose an external magnetic field B_z that is oriented along the z-axis, where the magnetic moments can align parallel or anti-parallel. The following component of MRI is a Radio Frequency (RF) impulse Deserno (2010); this impulse pushes the proton down with a frequency equal to the rotation rate of the proton (always at the same point of the movement), making it turn horizontally and parallel to the ground. After RF excitement, an exponential relaxation, the proton restores to its equilibrium state Nordenskjöld (2014). The stored energy is released as a
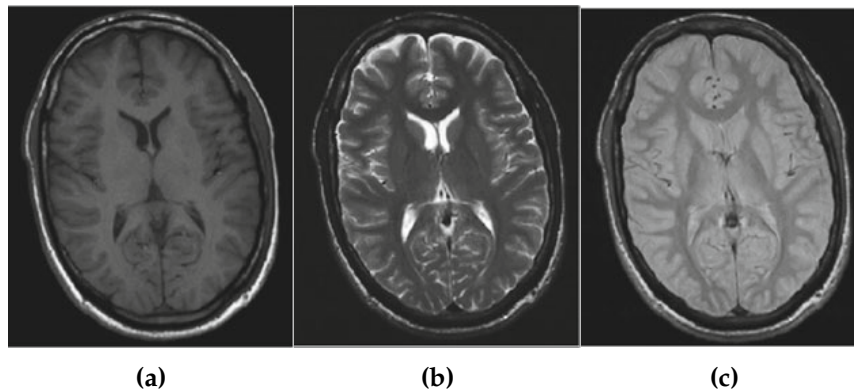
Figure 2.1: Images collected using different T_R and T_E combinations. a) $T1_w$, b) $T2_w$ and c) PD_w. Images taken from Deserno (2010).

signal, which can be detected and transformed into an image. Two independent effects control the relaxation process:

1. T2 relaxation (*spin–spin*).- it affects the phase of the spins. For water-based and fat-based tissues, T2 is in the $40 - 200$ ms and $10 - 100$ ms range, respectively.

2. T1 relaxation (*spin–lattice*).- it affects the parallel vs antiparallel alignment of spins. For water-based and fat-based tissues, $T1_w$ is in the $400 - 1200$ ms and $100 - 150$ s range, respectively.

The repetition time (T_R) denotes the rate of re-applying a sequence of RF, and the echo time (T_E) is the period between a transmission and data collection Nordenskjöld (2014), are set to produce differently weighted MRI sequences (Figure 2.1). T1-weighted ($T1_w$) MRI sequences are obtained if ($T_E \leqslant T_2$ and $T_R \approx T_1$), T2-weighted ($T2_w$) MRI sequences are produced if ($T_E \approx T_2$ and $T_R \geqslant T_1$) and proton density-weighted PD_w MRI sequences are got if ($T_E \leqslant T_2$ and $T_R \geqslant T_1$) Deserno (2010). $T1_w$ MRI sequences include increased anatomic detail relative to $T2_w$ MRI sequences, $T2_w$ MRI sequence is better for assessing edema and has generally shorter imaging times and PD_w is an intermediate MRI sequence, which seeks to combine $T1_w$ and $T2_w$ MRI sequences characteristics El-Dahshan et al. (2014). Meanwhile, certain tissues will show up as high signal intensity on grayscale imaging and other as low

signal intensity. On $T1_w$ images, fluid in the tissues presents as intermediate to low signal intensity, and fat as high signal intensity. On the other hand, fluid on $T2_w$ appears as a high signal and fat as a low signal. Another MRI sequence is the Fluid Attenuated Inverse Recovery (FLAIR); this modality presents a high contrast between tissue and lesions Valverde et al. (2017).

2.2 Segmenting brain tissues in MRI scans

An image is a collection of measurements arranged in a two-dimensional (2D) or three-dimensional (3D) space. These measurements are obtained in medical images from different image acquisition techniques, such as X-ray imaging, ultrasound imaging (sonography), or MRI. An image is called *scalar* if a single measurement is made at each of its points; if more than one measurement is made, the image is referenced as a multi-channel image. These images may be acquired in the continuous domain (X-ray film) or discrete space (MRI scans). In 2D discrete images, the location of each measurement is called a pixel. On the other hand, in a 3D image, it is called a voxel Pham et al. (2000).

Let the image domain be given by Ω; then the segmentation problem is to define the sets $R_k \in \Omega$, whose union is the entire domain Ω Gonzalez et al. (2004). Hence, the sets that make up a segmentation must satisfy:

$$\Omega = \bigcup_{k=1}^{K} R_k \tag{2.1}$$

where $R_k \cap R_j = \emptyset$ for $k \neq j$. When the regions R_k are not connected, the process is called partition, and the sets themselves are called classes.

2.2.1 Brain tissue segmentation

Image Segmentation refers to the action of parcellate an image into a set of homogeneous and non-overlapping objects that share a common set of attributes Despotović et al. (2015). The segmentation result is an image of labels identifying each object in the image. In the case of brain MRI scans, brain tissue segmentation refers to the parcellation of the intracranial tissues from a brain MRI scan into specific tissue types El-Dahshan et al. (2014); Despotović et al.
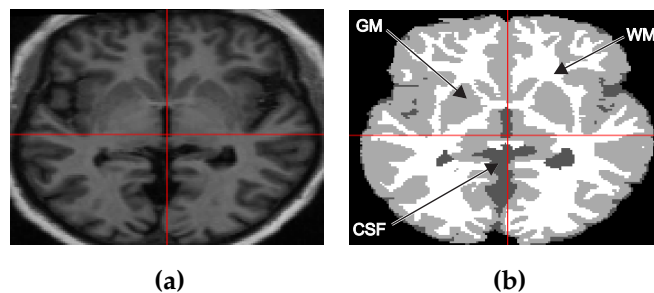
Figure 2.2: An example of the brain MRI segmentation: (a) Raw brain $T1_w$ MRI scan; (b) the segmented image with three labels WM, GM, and CSF.

(2015). In this research, the brain tissue segmentation separates the voxels into three main tissue types: CSF, GM, and WM.

The CSF is not a tissue but a liquid that surrounds and protects the brain acting as a shock absorbent. It also works as an agent for nutrition transportation across the central nerve system Nordenskjöld (2014). The GM represents the cerebral cortex, and the WM represents connecting fibers between neurons, allowing communication between different brain regions. The GM surrounds the brain, but it also has some parts embedded in the WM Deserno (2010). Images of brain MRI scans depict CSF class in dark-gray, GM appears gray-brown, and WM forms a white-colored area. Figure 2.2 shows an example.

2.2.2 Pre-processing: Skull-stripping and bias field correction

Before the segmentation process, it is convenient to apply different pre-processing procedures. The most common are skull-stripping and intensity correction. It is convenient for brain tissue segmentation to remove non-brain tissues (fat, skull, dura, and marrow) whose intensities overlap with the brain tissues' intensities of interest. This task is known as skull-stripping. The Brain Surface Extractor (BSE) Shattuck et al. (2001) and Brain Extraction Tool (BET) Smith (2002) are two of the most effective algorithms to achieve skull-stripping.

BET is an automatic method for segmenting brain and non-brain tissues on MRI sequences. The BET models the gray-level intensity of brain/non-brain tissues, estimates the center of gravity of the head, defines an initial sphere based on the center of gravity, and deforms the tessellated sphere outward the
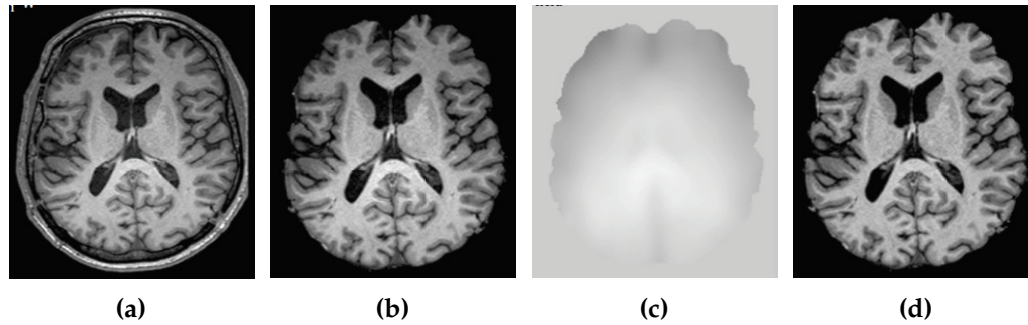
Figure 2.3: Pre-processing steps: (a) raw brain MRI scan in $T1_w$ MRI sequence; (b) brain MRI scan after skull-stripping; (c) the bias field; (d) brain MRI scan after bias field correction. Images taken from Despotović et al. (2015).

brain surface.

BSE is an edge-based method that employs anisotropic diffusion filtering to separate the brain from non-brain tissues. The BSE rationale relies on the fact that the brain is the largest area surrounded by a strong edge. Its work principle can be summarized in four phases:

1. Filtering the image to remove irregularities and enhance the image contrast.

2. Detecting edges in the image.

3. Performing morphological erosions and brain isolation.

4. Performing surface clean up and image masking.

Sometimes, it is required to execute more than once the skull-stripping algorithm because non-brain material remains on the MRI after one iteration. Intensity inhomogeneity (INU) is depicted as a shading effect (intensity bias) exhibited across an image Makropoulos et al. (2018) (see Figure 2.3). INU artifacts degrade the segmentation performance of methods that assume that a tissue class's intensity value is constant over the image. INU correction, also called bias field correction, is commonly performed with the N3 Sled et al. (1998) and N4 Tustison et al. (2010) algorithms. Figure 2.3 depicts an example of images obtained from the pre-processing procedures on a brain MRI scan.

2.2.3 Partial Volume Effect

The PVE causes a digital blurring in the image because the pixels/voxels'
gray value results from a mixture of signals from different objects. PVE is the
principal artifact that affects the segmentation process in the medical image
domain. PVE is produced by two phenomena: spatial resolution effect and
spatial sampling effect Soret et al. (2007). Spatial resolution refers to the
imaging system's property to differentiate the signal coming from two different
objects separated by a minimum distance Grupen and Buvat (2011). Spatial
resolution is measured in terms of the full width at half maximum (FWHM)
of the point spread function obtained by imaging a point source Grupen and
Buvat (2011). Therefore, any object with dimensions less than two to three
times the FWHM will be affected by the spatial resolution effect.

The sampling effect is known as the Tissue Fraction Effect (TFE) Grupen
and Buvat (2011); it is produced because the voxels do not match with the
borders of the objects of interest correctly. Therefore, the gray value of voxels
represents the mixture of signals from different tissue types. Figure 2.4, depicts
an example of TFE. The ROI, represented with a gray value of 8, is surrounded
by a background with a gray value of 1 (a). Sampling with high resolution
(b), taking the weighted average of the original gray values of the ROI and the
underestimation is lower (c). For low resolution (d), the ROI signal intensity
underestimation is large

Even with the highest spatial resolution, there would still be PVE caused
by image sampling for as long as resolution remains above atomic level Soret
et al. (2007). The TFE often makes difficult the correct definition of the borders
between the objects in the image. In this sense, a brain MRI scan contains two
types of voxels:

1. *Pure voxels*: voxels with a high probability of containing only one tissue
   type present.

2. *Mixels*: voxels with a high probability of containing a high level of mixture
   of two or more tissue types present.
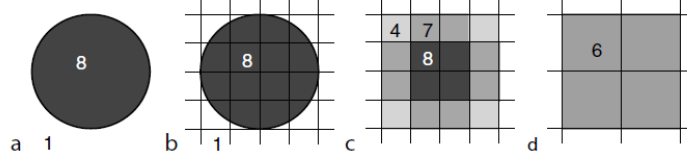
Figure 2.4: TFE resulting from spatial sampling. Image taken from Grupen and Buvat (2011).

2.2.4 Partial Volume Segmentation

Strictly speaking, all the pixels-voxels are mixed elements since the PVE is always present. Furthermore, the quantification of TFE is only possible on synthetic data. Hence, the segmentation methods only estimate the number of mixels in an image when a level of mixture is defined Bricq et al. (2008); Valverde et al. (2017). Therefore, most of the current methods that address the brain tissue segmentation problem are developed to determine each tissue type's relative fraction within each voxel. For this reason, this kind of segmentation is named Partial Volume Segmentation (PVS) Noe and Gee (2001); Van Leemput et al. (2003). Since partial segmentation results assessment is not possible in real data, it is necessary to transform partial labels into crisp. Those methods that consider partial labels in their procedures achieved better segmentation performance than methods that do not Valverde et al. (2015).

2.2.5 Dice similarity coefficient

The Dice similarity coefficient (DSC) is the most used metric for validating medical volume segmentation. The DSC is defined by:

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{2.2}$$

where TP is the number of true positives between the segmentation result from each method and the ground truth. FP corresponds to the number of false positives, and FN is the number of false negatives Taha and Hanbury (2015). We assessed the segmentation performance of our method on the CSF, GM, and WM classes individually.

2.2.6 Atlas

Brain atlases are labeled data that specify brain structures' location. The popularity of atlases stems from the guide that provides in the segmentation process in areas of poor contrast and helps distinguish between tissues of similar intensities Devi et al. (2015). The term atlas refers to a pair of images: atlas template image and atlas label image. The atlas template image depicts an MRI of the brain, while the atlas label image denotes tissues at every voxel Makropoulos et al. (2018). To avoid confusion, we use the term atlas to refer us to the atlas label image in this document. There are two types of atlases: single-subject and probabilistic. The single-subject atlases assign a single tissue label at each voxel. On the other hand, the probabilistic atlases define the structure probability of each structure/tissue at each voxel. The process of generating both types of atlas is different. Single-subject atlases are manually delineated, while probabilistic atlases are formed by averaging automatically derived segmentations Makropoulos et al. (2018). The use of a single-subject or a probabilistic atlas depends on the type of problem to be addressed. Single-subject atlases are useful for identifying brain substructures; probabilistic atlases effectively recognize large brain structures such as CSF, GM, and WM. In this research, we follow the probabilistic atlas approach and its application.

Montreal Neurological Institute (MNI) developed one of the most successful atlas used until now, named MNI152 linear Brett et al. (2002). This atlas was built with 152 MRI volumes from healthy people. The MNI152 linear was adopted to define standard anatomy by the International Consortium of Brain Mapping (ICBM). In 2009, the MNI released an updated version, the MNI152 NLIN 2009 Fonov et al. (2011). This atlas exhibits the best resolution and detail to date. It comes in three versions that each have symmetric and asymmetric subversion.

1. ICBM 2009a Nonlinear Symmetric.- atlas in $1.0 \times 1.0 \times 1.0$ mm^3 resolution for CSF, GM, and WM.

2. 2009b Nonlinear Symmetric.- atlas in $0.5 \times 0.5 \times 0.5$ mm^3 resolution for CSF, GM, and WM.

3. ICBM 2009c Nonlinear Symmetric.- atlas in $1.0 \times 1.0 \times 1.0$ mm^3 resolution for CSF, GM, and WM. Sampling is different from 2009a template.

2.2.7 Atlas co-registration

Atlas co-registration is the process used to transfer the labels of an atlas (source) to an unlabeled target subject (target) Makropoulos et al. (2018). In brain tissue segmentation, the prior tissue maps (CSF, GM, and WM) are co-registered to the target subject to provide spatial information to the segmentation model. Nevertheless, an appropriate atlas does not always exist for the data at hand. For example, Figure 2.5 depicts the Atlas co-registration output of the brain MRI scan, Figure 2.5(a), and the WM prior tissue Figure 2.5(b). As we can see in Figure 2.5(c), there is a considerable mismatch (in pink) for the WM. The Atlas quality and its correspondence with the target affect the estimated transformation Mayer and Greenspan (2009). BRAINSFit Johnson et al. (2007) is one of the most used methods to perform atlas co-registration. BRAINSFit algorithm uses the Mutual Registration algorithm Magnotta et al. (2002) to register a three-dimensional volume to a reference volume. However, its performance is compromised when it faces values with asymmetric voxel sizes.

(a) (b) (c)
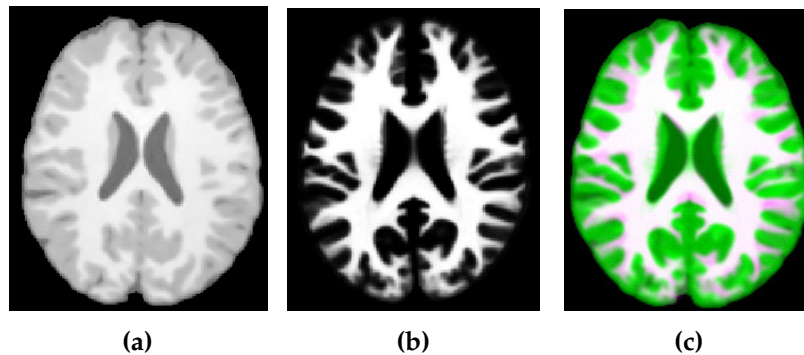
Figure 2.5: Example of atlas co-registration. (a) raw brain MRI scan in T1$_w$ sequence, (b) WM prior tissue, and (c) mismatch, marked in pink.

2.2.8 Root-mean-square error

The root-mean-square error (RMSE) is standard metric to asses the agreement between fractional volumes Bricq et al. (2008). The RMSE is defined by:

$$\text{RMSE} = \sqrt{\frac{1}{\Omega}(\mathbf{x}_{j_{test}} - \mathbf{x}_{j_{true}})^2} \tag{2.3}$$

where Ω is the intracranial mask (total voxels), $\mathbf{x}_{j_{test}}$ represents the estimated fraction of tissue for the j-th voxel and $\mathbf{x}_{j_{true}}$ represents the true fraction of tissue for the j-th.

2.2.9 Non-overlapped small regions for segmenting irregular objects

The tissue heterogeneity on the brain tissues, together with its irregular structure, increase the impact that the PVE projects on them. This issue is a challenging obstacle that the segmentation methods have to overcome. The use of non-overlapped small regions (subregions) is useful for segmenting irregular objects as brain tissues Al-Dmour and Al-Ani (2018); Kong et al. (2018). Each subregion is almost homogeneous in a feature space, while neighboring-regions possess different characters.

Watershed algorithm

The Watershed algorithm Meyer and Beucher (1990) is a region-based segmentation approach that produces distinct regions, even in low contrast
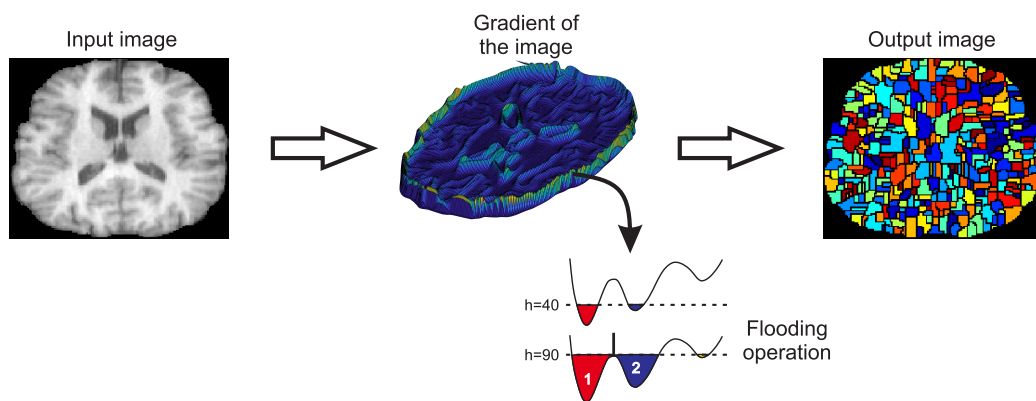
Figure 2.6: Visual example of the flooding operation performed by the watershed algorithm.

images Subudhi et al. (2018). We use the Watershed algorithm to oversegment a brain MRI scan into subregions to group them into the three major classes. The Watershed algorithm is a morphological gradient-based segmentation technique. The method considers the gradient image as a *topographic surface*, where different gradient values of the image correspond to different *heights*, forming catchment basins Subudhi et al. (2018). The algorithm then performs a flooding operation, starting from the regional minimum, filling the catchment basins. When two different bodies of water meet, a dam is built between them, forming watershed lines Saikumar et al. (2012). The process continues until the water level covers all the points on the map. The outcome of the algorithm is a labeled image with R small regions.

2.3   Machine learning definitions

2.3.1   Pattern recognition for brain tissue segmentation

Pattern recognition is defined as the process where a given pattern or signal is assigned to one of the prescribed classes Jain et al. (1999); Haykin et al. (2009). A pattern often is represented by a set of d scalar measurements x_i (features). This set of measurements is viewed as a d-dimensional feature vector $\mathbf{x} = (x_1, \cdots, x_d)$. Most of the current methods that address the brain tissue segmentation problem base their operation principle upon a set of patterns that describe the regions of interest. Those features based on texture analysis are particularly effective in the medical domain. This behavior is because some physical properties of biological samples (for example, color, density, or texture) are effectively described by texture features and provide enough information for the models to perform efficient segmentation Xu et al. (2004).

Haralick 2D texture features Haralick et al. (1973) are some of the most successful features used in pattern recognition in the medical domain El-Dahshan et al. (2014); Despotović et al. (2015). These features are computed from the gray-level intensity values on the image. Its calculation consists of the co-occurrence counts of the same gray-level intensity values in two pixels/voxels separated by an oriented separation vector within a sampling window. The counts are stored in a matrix from which a set of statistics are

calculated. A significant limitation with the 2D approach is that the window size around the pixel of interest should be relatively large to distinguish different regions correctly. If the sampling window size is too large, distant pixels influence features belonging to the pixel of interest. In other words, for bigger windows, features will be blurred.

Tesař et al. (2008) introduced the Haralick 3D texture features to tackle the limitations of the 2D approach. This approach adopts the fact that, in 3D space, the number of voxels around the diameter d of a given voxel (proportional to d^3) is larger than in the 2D space where the number of pixels is proportional to d^2. Therefore, the amount of texture data, which is in terms of the number of voxels or pixels in such surroundings, is better characterized for small volumetric objects.

2.3.2 Unsupervised learning

Unsupervised learning, also known as Clustering, has the goal of building decision boundaries based on unlabeled training data, where no prior tissue information is available Jain et al. (2000). Hence, clustering methodology discovers interrelations, making some assumptions about the data in the patterns, with minimum assistance from an external source. Clustering demands human intervention, but in this case, it occurs at the end of the process. The reliability and consistency with which a clustering algorithm can organize unlabeled data is a good reason to use it in brain tissue segmentation. Jain et al. (1999) defined the custom components of a clustering task:

- Pattern representation.- the number of classes, the type, scale, and number of available patterns to the clustering algorithm.

- Definition of pattern proximity.- similarity distance function defined on pairs of pattern

- Clustering or grouping.- the output grouping can be a partition of the data into non-overlapping groups (hard) or a vector, for each pattern, with the degree of membership for each of the output clusters (fuzzy)

- Assessment of output.- if the analyzed data set contains a ground truth, it is possible to assess the clustering method's performance quantitatively.

Most of the time, it is challenging, expensive, or even impossible to label each training sample with its true category reliably for the brain tissue segmentation problem. Then, there is not enough labeled data for training classification methods. Therefore, clustering methods are preferred to face the segmentation problem.

2.3.3 Self-Organizing Map

The Self-Organizing Map (SOM), proposed by Kohonen (1990), is a kind of Artificial Neural Network (ANN) composed, in its original formulation, of one input and one output layers. The increasing popularity of SOM to approximate pattern recognition problems lies in its low dependence on domain-specific knowledge (relative to model-based and rule-based approaches) and its flexibility to include a wide range of strategies for improving its performance.

The goal of SOM is to transform a set of input patterns of an arbitrary dimension into a discrete one or two-dimensional map. The input layer consists of a lattice of neurons $i, i = 1, ..., N$ each with a prototype (feature vector) $\mathbf{w}_j \in \mathbb{R}^n$ attached to it. These prototypes fit the data during the training process, and such adaptation is reflected in the neurons' topological sorting. Each output neuron represents a group prototype in the output layer, and new input patterns are mapped to the most similar prototype Haykin et al. (2009).

The algorithm responsible for forming SOM, which is unsupervised, begins with the initialization of the neurons' synaptic weights in the input layer. The weight initialization process can be accomplished by assigning small values, in an interval $[0 \quad 1]$, taken from a random number generator or using a linear initialization approach Su et al. (1999). After initializing the network, two fundamental processes, named activation and learning, are involved in forming the SOM Negnevitsky (2005).

Activation

Activate the SOM network by applying the input vector $\mathbf{x}$, and find the winner-takes-all, also known as Best Matching Unit (BMU), neuron $j_\mathbf{x}$ at iteration t, using the minimum-distance Euclidean criterion.

$$j_\mathbf{x}(t) = \underset{j}{\arg\min} \|\mathbf{x} - \mathbf{w}_j\| \tag{2.4}$$

where n is the number of neurons in the input layer, and m is the number of neurons in the output layer.

Learning

Update the synaptic weights as:

$$\mathbf{w}_j(t+1) = \mathbf{w}_j(t) + \Delta\mathbf{w}_j(t) \tag{2.5}$$

where $\Delta\mathbf{w}_{ij}(t)$ is the weight correction at iteration t. The weight correction is determined by the competitive learning rule

$$\Delta\mathbf{w}_j(t) = \begin{cases} \eta[\mathbf{x} - \mathbf{w}_j(t)], & j \in \Lambda_{jk}(t) \\ 0, & j \notin \Lambda_{jk}(t) \end{cases} \tag{2.6}$$

where η is the learning rate parameter, $\Lambda_{jk}(t)$ is the neighborhood function centered around the BMU_{j_x} at iteration t, and encompassing a set of excited (cooperating) neurons denoted by k Haykin et al. (2009). Let d_{jk} denote the lateral distance between the BMU_{j_x} and the excited neuron k. Then, assuming that the topological neighborhood Λ_{jk} is a unimodal function, it attains its maximum value at the BMU_{j_x} for which the distance $d_{jk} = 0$, while $\Lambda_{jk} \to 0$ when $d_{jk} \to \infty$; this is a necessary condition for convergence. The neighborhood can be defined as a square or hexagon. However, the smooth Gaussian kernel is mostly used Engelbrecht (2007). Figure 2.7 shows a diagram of a two-dimensional SOM, where each neuron, represented by a hexagon, is connected to the input vector. The distribution of neurons on the map follows a hexagonal pattern. The neighborhood function Λ_{jk}, depicted in a shading green color, covers the neurons around the BMU. The closer neurons, covered by a dark green, receive a more significant stimulus than the further ones.

Once the learning phase ends, the SOM provides an accurate approximation of the input data distribution coded in the prototypes, while the ordering on the map reflects their similarity relationships in the data space Merényi and Taylor (2019). In the mapping phase, a new input data is clustered with its BMU. Then, the nearby inputs are mapped to nearby BMUs.

SOM architecture

The two-dimensional architecture of the SOM network is adequate for approximating similar relationships of high-dimension data. Two-dimensional rectangular arrangements are frequently used for simplicity. However, hexagonal arrangements are more illustrative, accurate, and preferable Kohonen
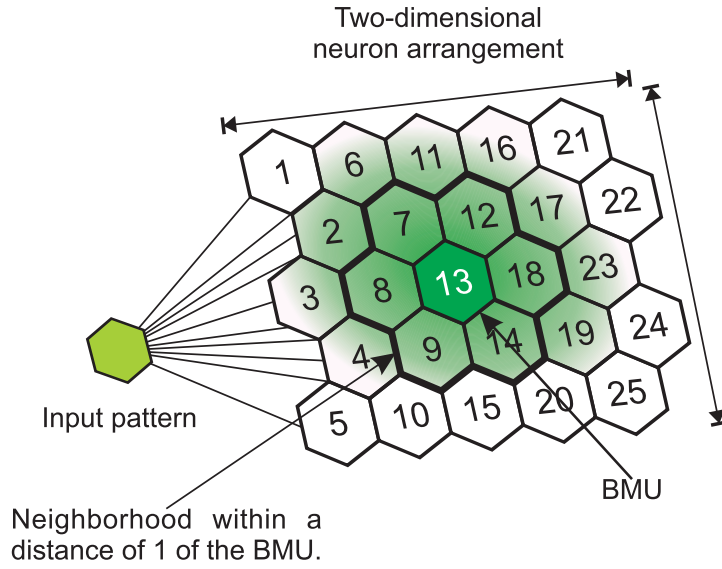
Figure 2.7: Hexagonal structure of a 5-row, 5-column SOM network.

(1990). The hexagonal topology effectively displays the map because the prototype neurons are distributed more smoothly than when using a rectangular topology Delgado (2017). Hence, we chose the hexagonal topology because it provides hexagonal neighborhoods and facilitates the maps' visual appreciation of possible patterns.

2.3.4 Fuzzy sets

Suppose X is the universe of discourse, and $x \in X$ is a specific element of the domain X. Then, the fuzzy set A of universe X is defined by a characteristic function (membership function) of set A:

$$\mu_A(x) : X \to [0, 1] \tag{2.7}$$

where

$\mu_A(x) = 1$ if x is totally in A;

$\mu_A(x) = 0$ if x is not in A;

$0 < \mu_A(x) < 1$ if x is partly in A;

This way, for any element x of universe X, the value of the membership function $\mu_A(x)$ represents the membership degree to which x is an element of set A. If the valuation set is restricted to $[0 \quad 1]$, A is called a crisp set. Fuzzy

sets theory has been used to deal with the uncertainty produced by the PVE in medical image segmentation Zhang et al. (2017); Prakash and Kumari (2017); Jiang et al. (2019), especially when clusters present severe overlap. In this research, membership levels are calculated using a univariate approach based on spatial models for fuzzy grouping Pham (2001b). This model facilitates the incorporation of penalty terms in the Fuzzy C-Means (FCM) objective function to constrain the membership functions' behavior.

2.4   Pseudo-lables

Pseudo-labels are target classes for unlabeled data as if they were true label Lee (2013). Instead of manually labeling the unlabelled data, the pseudo-labeling process approximates labels using the outcome from an initial clustering method Saito et al. (2017). The pseudo-labeling process follows the assumption that the voxels with high belonging to a particular cluster have a high possibility of being in the correct class Ahmadvand and Daliri (2015). This means that each $\mathbf{x}_j$ is linked with a pseudo-label $\widehat{y}_j$ as follows:

$$\mathbf{x} = [\mathbf{x}_j, \widehat{y}_j] \tag{2.8}$$

where $\widehat{y} = \text{argmax}_k \, \psi(\mathbf{x}_j)$. The value of $\psi(\mathbf{x})$ is the membership/probability level, computed by an initial clustering method. The pseudo-labels explicitly force the target class-conditional structure to be more discriminative Chen et al. (2019). Due to the lack of sufficient labeled data on brain tissue segmentation problem, pseudo labels represent a compelling candidate to support the training of methods that face problems with a scarce amount of labeled data.

2.5   Summary

The concepts of PVE, atlas, brain tissue segmentation were defined. The way brain segmentation will be carried out was defined, dividing each scan's voxels into three major classes CSF, GM, and WM. On the other hand, the machine learning methods mentioned were fuzzy sets, SOM, and pseudo-labels. These three approaches are the basis for developing a model to overcome the

challenges introduced by the PVE and incorporating prior knowledge in the brain tissue segmentation problem.

THE STATE OF THE ART

The brain MRI segmentation is not a trivial task because the problem is ill-posed, and the variety in the configuration of medical image devices produce MRI scans with different image properties and artifacts combinations. This diversity led to the development of a wide range of techniques for brain tissue segmentation. Figure 2.2 depicts an overview of the most popular methods for brain MRI segmentation. Note that there is no universal method suitable for all brain MRI scans, nor the existing methods have an equal performance for a particular type of brain MRI scans. For example, some of the methods use the gray-level histogram exclusively Thevenot et al. (2014), while some integrate spatial image information to be robust for noisy environments. Some methods use the gradient representation of the image to delineate the boundary of the ROIs Subudhi et al. (2018); Wu et al. (2016), while others separate large connected region of voxels of the image sharing similar intensities Grande-Barreto and Gómez-Gil (2018). On the other hand, the use of probabilistic Dong and Peng (2014) or fuzzy set-theoretic approaches Zhang et al. (2017) effectively circumvent the challenges introduced by the PVE and other image artifacts. Hybrid segmentation methods combine different complementary approaches into a combined procedure to overcome many of the disadvantages of each method alone and improve segmentation performance Shenoy et al. (2016); Kong et al. (2018); Rajchl et al. (2016). The most prominent distinction among the most successful models that have addressed brain tissue segmentation is the incorporation of atlas information and those that do not. Next, we briefly describe the most popular ones.

Brain MRI segmentation methods

Contour and shape based

These methods use boundary/edge information to segment the ROI.

Region based

These methods use statistical descriptrors in an energy minimization framework to perform the segmentation.

Statistical based

These methods labels voxels according to probability values, which are determined based on the intensity distribution of the image.

Machine learning based

These methods use high dimensional features to cluster/classify the image into different classes.

Hybrid based

These methods combine the strength of two or more aproaches. These methods are often more accurate and robust than single methods but their strucutre usually is more elaborated

- Deformable models (Rajendran and Dhanasekaran., 2012)
- Active contour (Wang et al., 2019)
- Level set (Dera et al.,2015)
- Atlas based (Wu et al., 2016)

- Seed region growing (Zanaty and Asaad, 2013)
- Watershed (Li et al., 2007)
- Thresholding (Thevenot et al., 2014)

- MRF (Ahmadvand and Daliri, 2015)
- GMM (Dong and Peng, 2014)
- EM (Asburner and Friston, 2005)

- ANN (Zhang and Jiang, 2014)
- Random forest (Pereira et al., 2016)
- SVM (Moeskops et al., 2015)
- FCM (Zhang et al., 2017)

- FCM + ANN (Ortiz et al., 2013)
- Level set + MRF (Shahvaran et al., 2012)
- Deformable models + MRF (Shenoy et al., 2016)
- Supervoxel + Graphs (Kong et al., 2018)
- Atlas + ANN (Rajchl et al. 2015)
- Atlas + FCM (Valverde et al., 2017)

ANN	Artificial Neural Networks
EM	Expectation-Maximization
FCM	Fuzzy c-Means
GMM	Gaussian Mixture Model
MRF	Markov Random Field
SVM	Support Vector Machine

Figure 3.1: Overview of the most commonly used segmentation methods for brain MRI segmentation. Image inspired from El-Dahshan et al. (2014)

3.1 Fundamentals on brain tissue segmentation

A fundamental component in the structural brain MRI analysis involves the partition of MRI data into specific tissue types. A standard approach, not the only one, to cover the classification task consists of segmenting the brain into three main tissue types: CSF, GM, and WM. An essential component in the development of segmentation methods is the type of data to be processed. The $T1_w$ scans are the most popular MRI sequences used in the brain segmentation because they are acquired with high-field MR scanners with a magnetic field strength of 1.5 T or 3 T Despotović et al. (2015). Hence, $T1_w$ scans provide the best scan resolution. However, evidence showed that adding additional MRI sequences with different contrast properties (e.g.,$T2_w$, and PD_w) can improve the brain tissue segmentation Maillard et al. (2008); Mayer and Greenspan (2009).

Bias field and TFE are the most important artifacts that negatively affect the brain tissue segmentation performance Balafar et al. (2010). The bias field is the product of several physical factors in the image device. Some of these factors are spatial inhomogeneity of the magnetic field, the reception coil's sensitivity, and the signal response from the interaction between the magnetic field and the human body Deserno (2010); Despotović et al. (2015). The Bias field causes a smooth signal intensity variation within tissue of the same physical properties. On the other hand, the TFE arises from the low spatial resolution of the image device. The TFE produces the loss of small tissue regions, which produce uncertainty in the correct definition of tissue borders Grupen and Buvat (2011). Segmentation methods employ spatial information from both the analyzed image and previous knowledge to overcome the outlined phenomena.

Due to the medical domain's image formation process, a voxel's gray-level intensity is strongly related to its neighbors' gray-level intensities. Hence, spatial knowledge from the image provides important information for brain tissue segmentation. Markov random field (MRF) theory is effective for modeling the local properties of an image Zhang et al. (2001). MRF models are popular on brain MRI segmentation methods because they can decrease misclassification errors due to image noise. Similar to MRF, Spatial models for fuzzy clustering incorporate spatial information in the form of a penalty term in the

FCM objective function to constrain the membership functions' behavior Pham (2001a).

The use of prior knowledge (probabilistic atlas) is often used to initialize and constrain the segmentation process. The main advantage of prior knowledge is the possibility to segment any brain structure available in the atlas without any additional cost Lötjönen et al. (2010). It is important to point out that atlas-based methods use prior knowledge in the spatial domain rather than feature space. Note that the reliability of the prior knowledge's information depends on the correspondence between the atlas and the target subject. Factors like the difference in voxel size between the source and the target directly affects the co-registration process. The above is a critical point to consider before using the available prior knowledge. Next, we present a study of the state-of-the-art focused on spatial information use.

3.2 Spatial context information extracted from the images

Two different lines of action have been explored on this approach: supervised and unsupervised. Supervised methods provide the best results when facing data affected by bias field and noise presence because of the inherent capability to fit the database characteristics Pereira et al. (2016). Kong et al. (2018) presented a supervised method based on super voxels and graph theory for segmenting brain tissues in $T1_w$ scans. Firstly, each scan is over-segmented into subregions (super voxels) using the simple linear iterative clustering (SLIC) algorithm Achanta et al. (2012). An undirected weighted graph $G = \{V, A\}$ is constructed for each scan. The nodes $V = \{v_1, v_2, \ldots, v_N\}$ of the graph are the super voxels, and A is the weighted adjacency matrix Sandryhaila and Moura (2013). Then, the graph algorithm propagates the nodes of known labels to predict the nodes of unknown labels. The model was validated on synthetic and real neuroimage datasets. The model's output is a labeled image, where the voxels are grouped into three major classes (CSF, GM, and WM). Nevertheless, the extended SLIC algorithm demands a strong human interaction because the number of super voxels for each scan to be segmented must be set manually.

Artificial neural networks (ANN) are some of the most popular methods used for brain tissue segmentation because they have self-learning capabilities that enable them to produce better results as more data becomes available. This property is very useful when ANNs have access to labeled examples to face noisy data. Amiri et al. (2017) developed a multiple-classifier system using three Multi-Layer Perceptron (MLP) classifiers. Each MLP is trained with different labeled sets and initialization weights. Instead of using intensity feature alone, each voxel is described by its feature vector composed of textural and geometric moment features. This model was validated on a real dataset, and the results achieved by this method are remarkable. However, the authors did not provide the guidelines for the initialization of the weights from the three MLPs or select the labeled sets. Therefore, the replication of their results is compromised. Demirhan and Güler (2011) presented a method based on SOM and Learning Vector Quantization (LVQ). The model performs a coarse training, executed by SOM, while LVQ is used for fine-tuning. This model was validated on real data, but it presents an important issue concerning the training partition; the user must select only ten scans to train the model. Therefore, the rater bias influences the performance of the model. Abdelsamea et al. (2015) presented a supervised hybrid model, called the Self-Organizing Active Contour (SOAC). The SOAC is based on the combination of SOMs and an extension of the level set algorithm. In the SOAC model, the user provides training examples belonging to the true foreground Ω^+ and true background Ω^- to train two SOMs, one for recognizing Ω^+ and other for Ω^-. Then, the topological information learned by the prototypes from both SOMs is integrated into the level set algorithm's energy function. For the classification phase, the user chooses a subset of voxels to initialize the level set function. Note that Ω^+ contains labeled examples from only one class while Ω^- contains the other classes. Therefore, the training and testing process must be repeated three times for each scan to parcellate the brain into CSF, GM, and WM classes. From the above, we can see that small partitions of training data taken from the image improves the model's learning and performance. However, a reliable mechanism for selecting training data is lacking.

Unsupervised methods circumvent the necessity of labeled samples in the training phase by making assumptions about the data distribution to find

patterns in the data. Dong and Peng (2014) presented a model based on the combination of Local Gaussian Mixture Models (LGMM) Liu and Zhang (2013) and Non-Local Regularized FCM (NLRFCM) Caldairou et al. (2009). The NLRFCM is used to model the bias field and noise from small local neighborhoods in the image. The LGMM models each tissue's distribution by a linear combination of Gaussian function more accurately since NLRFCM reduces the negative influence of inhomogeneity and noise. The outlined approach achieves smooth segmentation results and preserves fine structures in brain images on $T1_w$ scans with high noise presence. However, the scope of this method is limited by the low spatial resolution. Sulcal CSF, which present thickness of one or two voxels.

Fuzzy approaches estimate the class of the voxels using a membership function Verma et al. (2014); Ji and Sun (2017). This property is particularly useful to tackle the uncertainty derived from PVE. A fuzzy clustering can be transformed to a hard partition in general manners, ergo using a maximum membership criterion Ortiz et al. (2013b). The FCM algorithm is the most popular fuzzy method to address brain tissue segmentation. FCM can be extended to incorporate multi-channel data Lin et al. (2012), histogram features Zhang et al. (2017) spatial context Zhang et al. (2014), modeling bias Pham (2001a), and combinations with a wide range of methods Choudhry and Kapoor (2016); Prakash and Kumari (2017). Pham (2001a) proposed the Spatial Models for fuzzy clustering, which comprise spatial context information, applicability to multi-channel data, uncertainty modeling, robust behavior against noise, and a straightforward implementation. Spatial Models introduced a penalty term that minimizes its value when the membership level for a particular class is large and the membership levels for different classes at neighboring pixels are small (and vice versa). In other words, spatial models establish a negative correlation between the membership level of a class and the membership values of different classes at neighboring pixels. Like most FCM-based extensions, the model's segmentation performance can improve if its initialization does not follow a random strategy.

Other unsupervised approaches for brain tissue segmentation are SOM-based methods. These methods can learn a model directly from the data, independently of any well-description, due to their cooperative learning strat-

egy. For example, Ortiz et al. (2013a) presented two methods based on SOM, named Histogram based Fast Segmentation (HFS-SOM) and Entropy-Gradient Segmentation (EGS-SOM). The first method employs histogram information extracted from the whole volume to train a model. The second method uses first and second-order features extracted derived from the spatial relationship among different voxels. Both methods were assessed on real $T1_w$ scans, but EGS-SOM performed better than HFS-SOM when facing different voxel resolutions. The SCSOTM Zhang and Jiang (2014) assumes that those voxels clustered into the same neuron came from the same localized region on the brain scan. However, that study does not provide enough evidence to guarantee that this condition always happens. To the best of our knowledge, the current SOM-based methods base discard the topological information generated by the learned model. This condition is a knowledge gap that has not been studied and can lead to a new paradigm.

3.3 Spatial information extracted from an atlas

Probabilistic atlases are labeled data that specify the brain's different structures and provide a probabilistic estimate of each structure at every voxel Makropoulos et al. (2018). This information is essential for segmentation approaches that require a prior probability model for each label. The atlas co-registration process is necessary to incorporate probabilistic atlases in the form of prior knowledge of the brain anatomy into the segmentation method. This process provides a one-to-one transformation that maps a pre-segmented atlas image to the target image that requires segmenting. Because the atlas is inherently segmented, all structural information is transferred to the target image Pham et al. (2000). The human brain's standard probabilistic atlas consists of three tissue probability maps for CSF, GM, and WM. The information provided by the atlas can be used to initialize and constrain the segmentation process, remove non-brain tissues, bound the PVE artifacts and increase the robustness and accuracy of a segmentation method Despotović et al. (2015).

3.3.1   Label propagation

Label propagation is the easiest and fastest way to assign a label to each input image voxel. This segmentation procedure relies on a registration process that aims to estimate the anatomical differences between a single-subject atlas and the input image volume. Global rigid and affine transformations are usually enough to achieve the task. It is assumed that the atlas is close to the subject's anatomy. However, substantial anatomical differences between the atlas and the input image volumes produce large registration errors causing critical segmentation errors. In order to overcome such errors, the atlas and/or the target volumes are split into smaller sub-volumes. Then, these sub-volumes are registered hierarchically using rigid and affine transformations Cabezas et al. (2011).

3.3.2   Multi-atlas segmentation

In Multi-atlas segmentation approaches, multiple atlas are co-registered with the subject target to better deal with the registration errors obtained from a single-subject atlas and better account for anatomical variability. In the standard procedure, a subset of atlases can be selected randomly or all the available atlases. A rate of 15 to 25 atlases is the number suggested to use when dealing with the segmentation of objects with well-defined shapes that can present slight deformations between images Aljabar et al. (2009). However, that number depends on the anatomical variability among the atlases and their correspondence with the subject Wu et al. (2016). Then, the labels from the prior tissue knowledge are propagated to the target subject using different strategies such as the majority voting from all warped labels predicting each voxel, graph cuts Wolz et al. (2009) or Bayesian frameworks Han and Fischl (2007). Multi-atlas segmentation approaches produce the best precision from proven algorithms for segmenting subcortical structures, but it demands a high computational cost in time and memory. Therefore, it is infrequent that this method is used to segment more general brain tissues such as CSF, GM, and WM.

3.3.3 Probabilistic atlases as prior knowledge

Conceptually, the incorporation of prior information is similar to the labeled examples in supervised methods, except that they are implemented in the spatial domain rather than in the feature space. However, both domains can be complemented each other. The prior information provides an estimation of the voxels' probability/membership to each class. Furthermore, prior knowledge can be used to reject or learn outliers that are not present in the atlas Cabezas et al. (2011); Despotović et al. (2015).

The FMRIB's Automated Segmentation Tool (FAST) Zhang et al. (2001) and Unified Segmentation (US) Ashburner and Friston (2005) are well-established methods fully dedicated to brain tissue analysis. The FAST method is based on the combination of MRF Li (2009), Expectation-Maximization (EM) algorithms, and prior knowledge Moon (1996). The US method is based on GMM and iteratively incorporates atlas co-registration and bias field correction at the same time. Both methods work with $T1_w$ MRI sequences, but they also handle a multi-channel data from $T1_w$, $T2_w$, and PD_w MRI scans.Partial Volume Classifier (PVC) Shattuck et al. (2001) is another well-established method based on the Maximum-A-Posteriori (MAP) classifier and spatial prior information of the brain. Warfield et al. (2000) developed a segmentation algorithm that iterates between a classification step to identify tissues and an elastic matching step to align a normal brain anatomy template with the classified tissues. The elastic matching step provides the segmentation by registering an anatomical atlas to a patient scan. This method presents two interesting characteristics. It supports $T1_w$, $T2_w$, and PD_w MRI scans to detect and reject outliers such as white matter lesions. Segmentation of cortical and subcortical classification grey matter from MR of neonates. However, the performance of the method is conditioned by the contrast among the tissues.

Ahmadvand and Daliri (2015) presented a hybrid model based on GMM and FCM algorithms to decrease the MRF modeling run-time for segmenting brain tissues on $T1_w$ MRI sequences. First, the model selects a subset of voxels with a high membership level to a particular cluster, assuming that those voxels have a high possibility of being in the right segment. The hyperparameters of MRF are computed using the initial segmentation. Then, the MRF method is

applied over the remaining voxels to complete the segmentation.

Probabilistic atlases can also be incorporated in fuzzy approaches, such as extensions of the FCM algorithm. The prior information is included, as a penalty, in the objective function to constrain the behavior of the membership functions Shiee et al. (2010). This penalty discourages unlikely configurations, such as a high membership value surrounded by high values of different classes. Probabilistic atlases also provide an alternative to replace random initialization on fuzzy approaches Valverde et al. (2017), which is one of the principal limitations of this kind of method. Furthermore, the output of these kinds of methods can improve the original probabilistic estimation of the atlas at every voxel. The latter is one of the main contributions of this type of method and is a fundamental piece for other types of studies Girard et al. (2014).

The incorporation of prior information from probabilistic atlases into ANNs architecture is another explored alternative. Rajchl et al. (2016) presented a hybrid method that used SOMs to learn the Gaussian Mixture Model instead of using the EM algorithm. In work proposed by Vrooman et al. (2007), a conventional k-Nearest-Neighbor (kNN) algorithm was trained with multi-channel data ($T1_w$, $T2_w$, and PD_w) and atlas prior information. Firstly, two trained physicians manually segmented twelve brain MRI volumes from healthy subjects. A generic atlas is then co-registered to each manual segmentation to generate the initial prior information for CSF, GM, and WM. Next, a set of training voxels are selected automatically using a pre-defined threshold over the enhanced prior probability tissue maps. The kNN classifier trains with the chosen samples, while the rest of the voxels are classified using the learned model. The main finding of this work is the automatic selection of training samples in the training phase. Larger threshold values (`threshold` > 0.8) lead to less available but more reliable tissue samples. On the other hand, lower threshold values (`threshold` < 0.5) favor the correct CSF classification, but it leads to burden computation times and memory usage. Therefore, the authors suggest using values $0.5 <$ `threshold` < 0.8 for those methods that choose training samples from prior probability information. Finally, in Table 3.1, we summarize the main advantages and drawbacks of each strategy outlined in this chapter.

Table 3.1: Advantages and Challenges of the different reviewed brain tissue segmentation strategies.

Approach	Description	Strengths	Challenges
Without atlas			
Unsupervised	Assumptions about the data distribution to find patterns in the data	Effective when there is no enough clinical experts neither labeled samples are available	Troubles facing data severely affected by image artifacts
Supervised	The model is provied with a subset of labeled samples to learn the parameters of each class	Effective facing data severely affected by bias field and noise presence	Reproducibility compromised because of bias observer immersed in different training data sets
With atlas			
Label propagation	Atlas labels are directly applied in image space	Intuitive Straigh-forward A single registration	No anatomical variability Atlas dependant Registration dependant
Multi-atlas	Multiple labels are combined in image space	Anatomical variability Outlier minimisation	Atlas selection and combination Multiple registrations
Prior knowledge	Atlas values are used in a probabilistic framework	A single registration Multiple input features	Atlas weighting Model estimation

3.4 Discussion

In this section, we proposed a dichotomy of the strategies for addressing the brain tissue segmentation. We proposed a separation between methods that incorporate prior tissue knowledge and those that do not. Because brain tissue segmentation is an ill-posed problem, there is a wide range of methods to provide the best approximation. However, no computational method provides superior performance for all cases; even manual segmentation varies its results due to observer bias. The current methods are developed around $T1_w$ scans, but $T2_w$ and PD_w provide information to improve the models' discriminant power. The inclusion of spatial information limits the image artifacts effect (bias field, noise, and TFE) in the segmentation process. A rising approach is the use of small partitions of the scans to train or initialize the model. This approach considers the isometry of the voxels; hence, the methods can be applied to MRI sequences with different resolutions.

Table 3.2 describes the characteristics and segmentation performance of

the methods mentioned in this chapter. Results from FAST until SOM were taken from the study reported in Valverde et al. (2015). It is essential to point out that the segmentation results from Table 3.2 are just for orientation since the databases are different. Even methods that have worked with the same database, they used different partitions. The majority of the reviewed methods use the DSC coefficient to evaluate the accuracy of methods. However, other measures, such as the Sensitivity and Specificity, False Positive Fraction, False Negative Fraction, and Jaccard index, can be used Taha and Hanbury (2015).

Table 3.2: Performance of some the most popular state-of-the-art methods.

| Author | Dataset | MRI sequence | Whole volume | DSC | | |
				CSF	GM	WM
Kong et al. (2018)	BrainWeb	$T1_w$	Yes	0.90 ± 0.01	0.92 ± 0.01	0.94 ± 0.01
	IBSR18	$T1_w$	Yes	0.57 ± 0.09	0.87 ± 0.03	0.85 ± 0.01
Amiri et al. (2017)	IBSR20	$T1_w$	Yes	0.42 ± 0.12	0.86 ± 0.03	0.81 ± 0.04
Demirhan and Güler (2011)	IBSR20	$T1_w$	60 scans per subject	-	0.66 ± 0.09	0.55 ± 0.12
Dong and Peng (2014)	IBSR20	$T1_w$	Yes	-	0.87 ± 0.02	0.86 ± 0.01
Zhang et al. (2001)	IBSR18	$T1_w$	Yes	0.12 ± 0.05	0.74 ± 0.04	0.89 ± 0.02
	IBSR20	$T1_w$	Yes	0.13 ± 0.04	0.68 ± 0.06	0.79 ± 0.10
Ashburner and Friston (2005)	IBSR18	$T1_w$	Yes	0.17 ± 0.08	0.81 ± 0.02	0.88 ± 0.01
	IBSR20	$T1_w$	Yes	0.21 ± 0.07	0.78 ± 0.06	0.81 ± 0.08
Shattuck et al. (2001)	IBSR18	$T1_w$	Yes	0.13 ± 0.06	0.70 ± 0.08	0.83 ± 0.07
	IBSR20	$T1_w$	Yes	0.13 ± 0.05	0.66 ± 0.11	0.63 ± 0.23
Pham (2001b)	IBSR18	$T1_w$	Yes	0.11 ± 0.06	0.70 ± 0.06	0.88 ± 0.03
	IBSR20	$T1_w$	Yes	0.14 ± 0.05	0.69 ± 0.09	0.77 ± 0.14
Pham (2001a)	IBSR18	$T1_w$	Yes	0.11 ± 0.06	0.71 ± 0.06	0.88 ± 0.03
	IBSR20	$T1_w$	Yes	0.15 ± 0.06	0.70 ± 0.10	0.77 ± 0.14
Kohonen (1990)	IBSR18	$T1_w$	Yes	0.11 ± 0.06	0.70 ± 0.07	0.87 ± 0.03
	IBSR20	$T1_w$	Yes	0.15 ± 0.06	0.69 ± 0.09	0.77 ± 0.14
Zhang and Jiang (2014)	IBSR20	$T1_w$	1 scan per subject	-	0.70 ± 0.04	0.63 ± 0.05
Ahmadvand and Daliri (2015)	BrainWeb	$T1_w$	10 scans per subject	-	0.88 ± 0.06	0.89 ± 0.02
Ortiz et al. (2013b)	IBSR18	$T1_w$	Yes	-	0.91 ± 0.03	0.92 ± 0.02
Valverde et al. (2017)	MRBrainS	$T1_w$ and FLAIR	Yes	0.81 ± 0.03	0.82 ± 0.02	0.87 ± 0.03
Rajchl et al. (2016)	MRBrainS	$T1_w$, $T2_w$, and FLAIR	Yes	0.81 ± 0.04	0.84 ± 0.01	0.88 ± 0.01

$T1_w$ from BrainWeb were acquired on a 1.5 T scanner with isotropic voxel size $1.0 \times 1.0 \times 1.0$ mm^3.

$T1_w$ from IBSR18 were collected using two 1.5 T scanners with three different anisotropic voxel sizes ($0.84 \times 0.84 \times 1.5$ mm^3, $0.94 \times 0.94 \times 1.5$mm^3, and $1.0 \times 1.0 \times 1.5$mm^3).

$T1_w$ from IBSR20 were collected using a 3.0 T scanner with anisotropic voxel size $1.0 \times 1.0 \times 3.0$ mm^3.

$T1_w$ from MRBrainS were collected using a 3.0 T scanner with isotropic voxel size $1.0 \times 1.0 \times 1.0$ mm^3.

$T2_w$ and FLAIR from MRBrainS were collected using a 3.0 T scanner with anisotropic voxel size $1.0 \times 1.0 \times 3.0$ mm^3.

3.5   Summary

In this chapter, we introduced the fundamentals of brain tissue segmentation. We presented a study of the state-of-the-art based on methods that incorporate spatial information in their architecture. The study's main contribution is the unveiling of the gap in knowledge in current methods that have addressed the brain tissue segmentation. The study showed some of the main databases used to validate the methods and compare the performance methods.

THE PROPOSED METHOD

The proposed method consists of two phases: adjusting a generic atlas to a specific target and segmentation of brain tissues on a brain MRI slices. The Gardens2 algorithm covers the first section, while the PLA-SOM meets the second one. We developed both methods to work together, but both algorithms can be used independently. The output from Gardens2 provides spatial constraints (prior tissue knowledge) to other segmentation methods. On the other hand, PLA-SOM can be applied to the output from an initial segmentation to improve its performance. Figure 4.1 depicts the pipeline of the proposed method.

4.1 Gardens algorithm

Initially, we proposed an algorithm named Gardens Grande-Barreto and Gómez-Gil (2018) for segmenting brain tissues. In the Gardens algorithm, a generic atlas supports the model's initialization, while the clustering was based on the feature space. The Gardens algorithm's output is a hard labeling; the selection of the atlas is a direct factor in the algorithm performance.

4.2 Gardens2 algorithm

We present a novel unsupervised algorithm named Gardens2. Instead of performing an exhaustive search for the correct atlas, the Gardens2 algorithm takes the output from a co-registration process and fits it to the target subject through a fuzzy function.

Gardens2 is composed of three main sections named Preprocessing, Fitting,
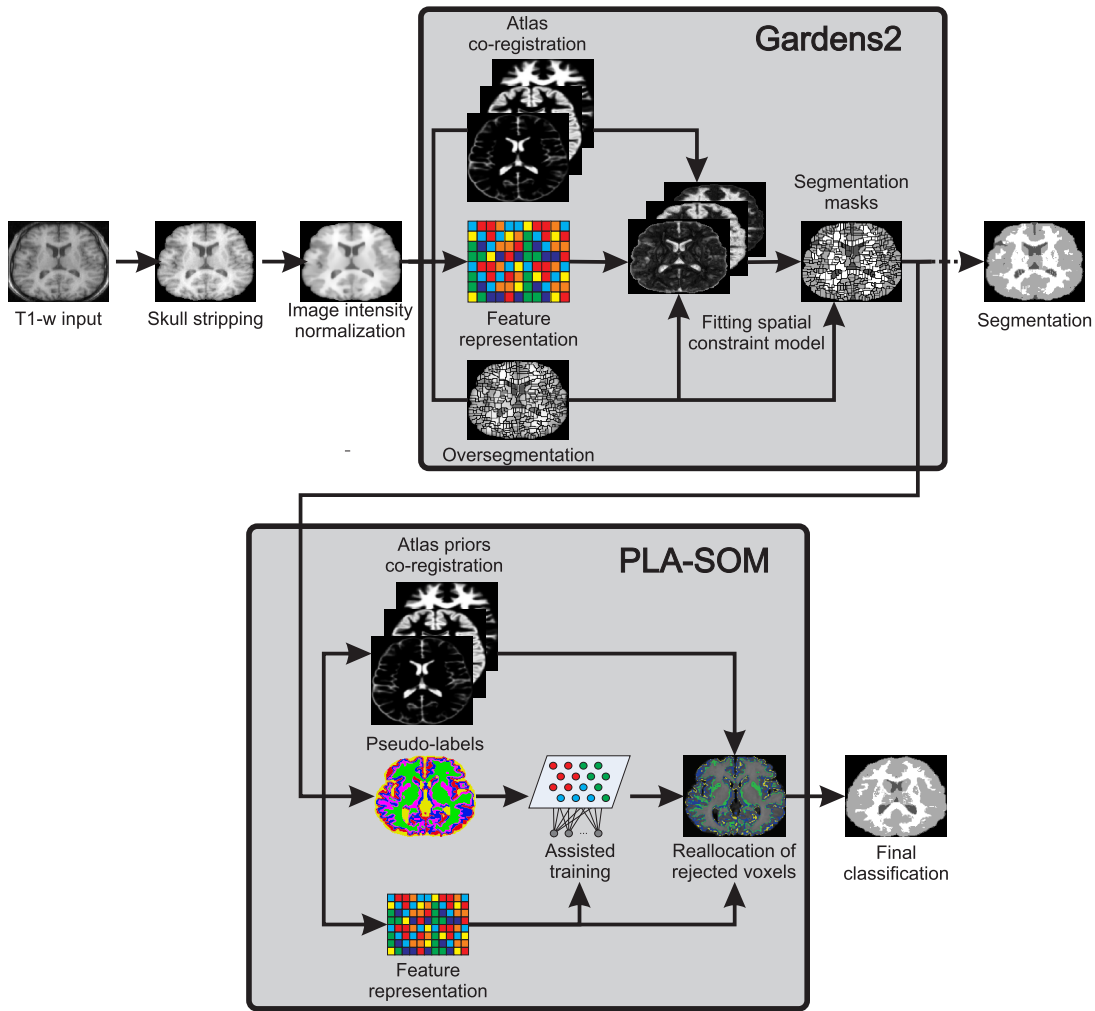
Figure 4.1: Pipeline of the proposed method.

and Tissue segmentation mask. The Preprocessing consists of skull-stripping, image normalization, feature representation, and atlas co-registration operations. In the Fitting section, a generic atlas is adjusted to a specific target using a novel fuzzy function. If it is required, Gardens2 can perform a hard segmentation using a segmentation mask described in the Section 4.2.7.

4.2.1   Skull stripping

In a raw brain MRI volume, some tissues interferes with the brain tissue segmentation. These non-brain structures are not of interest to the study of brain tissues presented in this manuscript; therefore, we removed them using

the BSE algorithm. We executed the BSE algorithm twice on each subject because non-brain material remains on the MRI after one iteration.

4.2.2 Image intensity normalization

Image intensity normalization is a standard procedure when working with MRI volumes. The adult brain's MRI intensity distribution lies in a narrow interval with intensity overlaps among tissue classes Despotović et al. (2015). This condition affects the performance of automatic methods for segmenting thin structures such as sulcal CSF. We performed a linear image intensity normalization on each brain MRI volume to enhance brain tissues' contrast. The linear normalization is performed as follows:

$$I_n = (I_o - c)\frac{b - a}{d - c} + a \tag{4.1}$$

where I_n is the normalized image, I_o is the original image, a and b are the target minimum and maximum gray levels, respectively. c and d represent the original minimum and maximum gray levels from I_o. For the experiments reported in this study, we set $a = 0$ and $b = 2^n$, with $n = 8$.

4.2.3 Feature representation

There are many feature representation approaches around brain segmentation. These approaches cover gray-level, histogram, wavelets, and texture analysis. The study in Ortiz et al. (2013b) showed evidence that the 3D Haralick texture features are one of the most useful features to train a model to separate brain tissues correctly. The reason for this result relies on the fact that 3D Haralick texture features can describe a relationship (gray-level intensity) of adjacent voxels in a 3D neighborhood, useful for analyzing brain tissues. In the proposed work, each voxel is described by a $15 - D$ feature vector x, which combines first and second-order descriptors extracted from a sliding cube that moves along the volume. First-order descriptors are the intensity level of the central voxel, mean, and variance of the voxels' intensity in the cube. On the other hand, we used texture features to describe the gray level intensity relationship of a set of voxels from a given region in the image. We used twelve textural descriptors computed from a 3D extension of Haralick texture features Tesař

Table 4.1: Set of features used to describe each voxel.

Index	Feature	Index	Feature	Index	Feature
1	Energy	7	Sum average	13	Voxel intensity
2	Entropy	8	Dissimilarity	14	Intensity mean
3	Correlation	9	Cluster shade	15	Intensity variance
4	Contrast	10	Cluster tendency		
5	Homogeneity	11	Maximum probability		
6	Variance	12	Difference variance		

et al. (2008) to represent the second-order descriptors. The extension uses cubes $(\omega \times \omega \times \omega)$ instead of square windows $(\omega \times \omega)$. These 3D textural features fit better than the original 2D to describe complex and compact structures like brain tissues because, in a 3D structure, a large number of voxels close to the central voxel are analyzed, and the influence of distant voxels is limited. Table 4.1 summarize the entire feature set.

4.2.4 Atlas co-registration

We chose the fuzzy tissue maps CSF, GM, and WM from ICBM 2009a Nonlinear Asymmetric Atlas template Fonov et al. (2011) to estimate the prior tissue model $\Gamma = \{\Gamma_{csf}, \Gamma_{gm}, \Gamma_{wm}\}$ for each subject. We used the BRAINS algorithm Johnson et al. (2007), with default parameters, to co-register the fuzzy tissue maps. The prior tissue model Γ will provide the spatial information for each voxel at every tissue.

4.2.5 Oversegmentation

We applied the Watershed segmentation algorithm on each brain MRI scan to oversegment them. The Watershed segmentation outcome is a labeled image $S = \{S_1, \ldots, S_R\}$, where R is the number of subregions in the image. Unlike other methods where the number of subregions (super voxels) is defined manually Al-Dmour and Al-Ani (2018); Kong et al. (2018), the Watershed algorithm defines R automatically, taking the surface covered by brain tissues on the image as reference. For example, Figure 4.2(a) depicts an input brain MRI scan on grayscale; Figure 4.2(b) shows the output of the Watershed algorithm. Gardens2 operates subregions S_i as basic units, except when a voxel by voxel

analysis is required for a precise evaluation.

4.2.6   Fitting spatial constraint model

A typical co-registration algorithm bases its work principle on the gray-intensity similarity. However, image artifacts on the MRI volumes and image properties of the atlas produce shifts on the co-registration output and yield undesired segmentation results. We assume that the correspondence between the atlas and the target can improve if the similarity measure is replaced by a novel one based on features that describe in more detail the gray-intensity relationships in the image (3D Haralick texture features) and a regularization term that imposes spatial constraints. We propose a novel fuzzy function to fit Γ with a specific MRI brain subject. The proposed function integrates feature and spatial domains to compute each voxel's membership degree on Γ as follows:

$$\widehat{\Gamma}_{jk} = \frac{(\|\mathbf{x}_j - \upsilon_k\|^2 + \gamma \sum_{l \in N_j} \sum_{m \in M_i} \Gamma_{lm}^q)^{-\frac{1}{(q-1)}}}{\sum_{j=1}^{C} (\|\mathbf{x}_j - \upsilon_k\|^2 + \gamma \sum_{l \in N_j} \sum_{m \in M_i} \Gamma_{lm}^q)^{-\frac{1}{(q-1)}}} \tag{4.2}$$

$\widehat{\Gamma}_{j,k}$ denotes the fitted tissue membership on the Atlas for a voxel j for a particular class k (see Figure 5.1). Variable x_j corresponds to the feature vector of j-th voxel. N_j is the set of neighbors voxels centered on j-th voxel, $M_i = \{1, ..., C\} \backslash \{k\}$ and $\{k \in C | C = \{csf, gm, wm\}\}$. γ is a scalar parameter

(a)   **(b)**   **(c)**
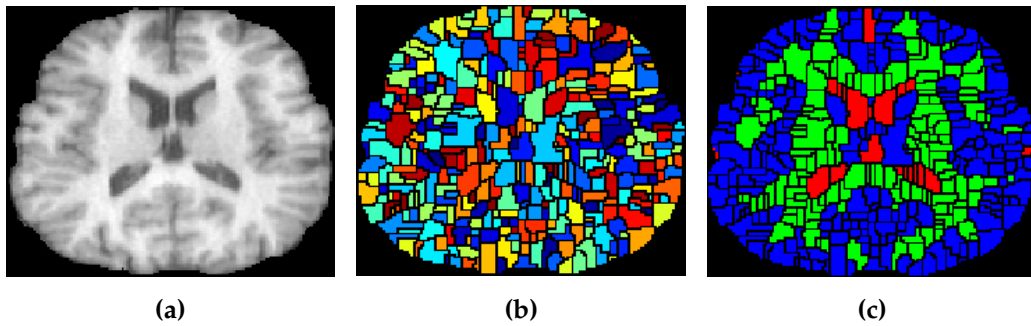
Figure 4.2: An example of brain tissue segmentation using Garden2. (a) Input image, (b) oversegmentation; each subregion S_i is drawn in a different false color. (c) Output image: segmentation of classes CSF (red), GM (blue), and WM (green).

that regulates the amount of prior tissue knowledge, which is used as a regularization term. q is a scalar parameter that controls the degree of fuzziness. υ_i represent the centroid of each class and is computed by:

$$\upsilon_i = \frac{\sum_{j=1}^{R_i} \sum_{k=1}^{V_j} x_{ijk_d}}{\sum_{j=1}^{R_i} V_j} \tag{4.3}$$

where x_{ijk_d} denotes d-th feature of the k-th voxel in the *j*-th subregion of the *i*-th cluster. V_j is the number of voxels in the *j*-th subregion. R_i is the number of regions in the *i*-th cluster; R_i is computed using the overlapping degree metric proposed in Grande-Barreto and Gómez-Gil (2018) as follows:

$$R_i = \frac{|S_j \wedge \Gamma_k|}{|S_j|} > 0.5 \tag{4.4}$$

where $|\cdot|$ denotes the cardinality of the set. The values of $\widehat{\Gamma}_{jk}$ lie $[0 \quad 1]$, and $\sum_{k=1}^{C} \widehat{\Gamma}_{jk} = 1$. $\widehat{\Gamma}$ can serves as a paradigm for automatic segmentation methods. Additionally, it can be tailored to perform a hard segmentation.

4.2.7 Tissue segmentation mask

Additionally to its primal task, the output of Gardens2 can be adapted to perform a hard segmentation. This secondary duty is achieved by transforming the results from Eq. (4.2) into a three-class tissue segmentation mask G_j, computed as follows:

$$G_j = \underset{k}{\operatorname{argmax}}(\widehat{\Gamma}_{jk}) \tag{4.5}$$

G_j is used to assign each subregion into a particular class using an overlapping criterion described as follows:

$$SEG_{S_r}(G_{csf}, G_{gm}, G_{wm}) = \begin{cases} CSF & \frac{1}{|S_r|}|S_r \wedge G_{csf}| > 0.5 \\ GM & \frac{1}{|S_r|}|S_r \wedge G_{gm}| > 0.5 \\ WM & \frac{1}{|S_r|}|S_r \wedge G_{wm}| > 0.5 \\ \ominus(S_r, k) & \text{otherwise} \end{cases} \tag{4.6}$$

where the operator $\ominus$ indicates a split operation for a subregion S_r. The split operation divides S_r on individual voxels that can merge with the *k*-th

---

Input : I_c, Γ, k

Output: Hard tissue segmentation

1 Transform I_c into a feature representation;

2 Compute S;

3 Generate prior-knowledge models Γ_{csf}, Γ_{gm}, and Γ_{wm};

4 Fit prior-knowledge models with (4.2);

5 Compute the three-class tissue segmentation mask G_j with (4.5);

6 Cluster S using (4.6);

Algorithm 1: Gardens2

cluster with the highest euclidean similarity. This split-and-merge strategy aims to correct the possible errors from the watershed segmentation. Algorithm 1 summarizes the entire process of Gardens2.

4.3   PLA-SOM

Our proposed algorithm starts by training a SOM model using the feature descriptors presented in Section 4.2.3 but complemented with pseudo-label information. We used pseudo-labels to support the inter-class separation and an intra-class compactness in training of SOM Cicek and Soatto (2019). Next, a novel fuzzy function uses the prototypes' data distribution and their ordering on the map. The proposed model also includes prior tissue knowledge, taken from Gardens2, to improve the final segmentation result.

4.3.1   Pseudo-labels

We alleviate the need for labeled data on the brain tissue segmentation problem using pseudo-labels computed from an initial clustering method. Let c denotes the number of classes $1 \leqslant c \leqslant n$, and define two sets of label vectors in $\mathbb{R}^c$ as follows:

$$\mathbf{N_{fcu}} = \left[y \in \mathbb{R}^c \middle| y_i \in [0,1]^c \forall i \right] = \text{(unconstrained) fuzzy/probabilistic,} \quad (4.7a)$$

$$\mathbf{N_{fc}} = \left[\boldsymbol{y} \in \mathbf{N_{fcu}} \Big| \sum_{i=1}^{c} y_i = 1 \right] = \text{(constrained) fuzzy/probabilistic,} \qquad (4.7b)$$

where N_{fc} is the convex hull, N_{fcu} is the unit hypercube in $\mathbb{R}^c$. Figure 4.3 depicts both sets for $c = 3$. For example, the vector $\boldsymbol{y} = (0.1, 0.6, 0.3)^T$ is a constrained label vector; its entries lie $[01]$, and $\sum_{i=1}^{c} y_i = 1$. The interpretation of $\boldsymbol{y}$ depends on its sources Bezdek et al. (1993). If $\boldsymbol{y}$ came from a fuzzy clustering method, $\boldsymbol{y}$ is a fuzzy label. If $\boldsymbol{y}$ came from a method such as a Gaussian mixture model, $\boldsymbol{y}$ is a probabilistic label. The cube $N_{fcu} = [0, 1]^3$ is called unconstrained label vector. Vectors such as $\boldsymbol{z} = (0.7, 0.2, 0.7)^T$ have each entry $[0 \quad 1]$, and $\sum_{i=1}^{c} z_i \neq 1$. In this research, we worked with constrained label vectors to address the voxel's belongings level at each tissue class.

We assume that voxels with a high belonging to a particular cluster are likely to be in the correct class Ahmadvand and Daliri (2015). We define the pseudo-labels $\widehat{y}$ as follows:

$$\widehat{y}(P_{csf}, P_{gm}, P_{wm}) = \begin{cases} \text{CSF} & P_{csf} > \tau_p & \& & P_{m \in M_i} < \tau_p \\ \text{GM} & P_{gm} > \tau_p & \& & P_{m \in M_i} < \tau_p \\ \text{WM} & P_{wm} > \tau_p & \& & P_{m \in M_i} < \tau_p \end{cases} \qquad (4.8)$$

where $\widehat{y}$ represent the pseudo-labels for classes CSF, GM, and WM. P is the level of belonging of each voxel to every class and $M_i = \{1, ..., C\} \backslash \{i\}$. In this study, the values from $\widehat{\Gamma}$ are used to compute the pseudo-labels. The scalar τ_p controls the level of belongings that a voxel must have to be chosen as pseudo-label. The selected voxels support the training phase, while the rejected ones get into a reallocation process.

Figure 4.4 depicts an example of the pseudo-label selection on a t-distributed Stochastic Neighbor Embedding (t-SNE) Maaten and Hinton (2008) on the BrainWeb data set. In Figure 4.4a, the ground truth y is depicted in red (CSF), blue (GM), and green (WM); in Figure 4.4b the pseudo-labels $\widehat{y}$ uses the same code color, the yellow points correspond to the voxels selected for the reallocation process.
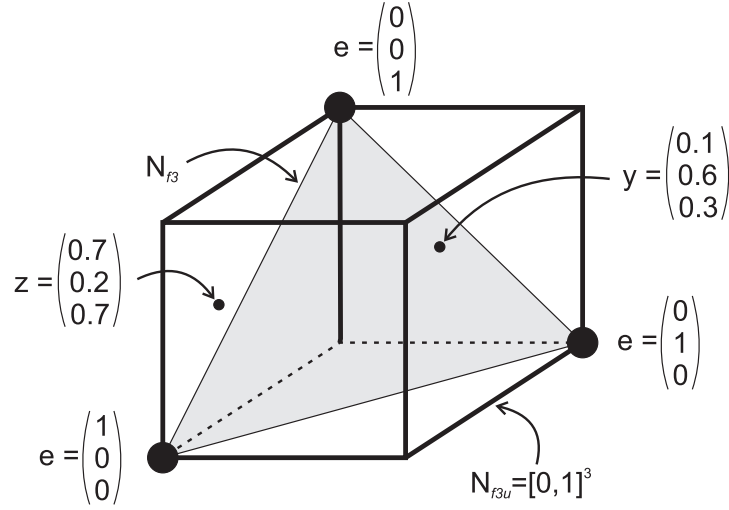
Figure 4.3: Example of constrained and unconstrained label vectors for $c = 3$. Image taken from Bezdek et al. (1993)

4.3.2 Assisted training of SOM using pseudo-labels

We intend with this strategy to assist the model in learning patterns from the same class and expel patterns from different ones, in order to improve the cluster separability. Let $\mathbf{x} \in \mathbf{X} | \mathbf{X} = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$ be a set of input patterns, and $\widehat{\mathbf{y}} \in \widehat{\mathbf{Y}}$ a class vector with the class information (pseudo-labels) for each $\mathbf{x}$. $\widehat{\mathbf{Y}}$ embeds the class information using unit vectors. For example $\mathbf{y} = [0, 1, 0]$, where a 1 in the component shows the class, and 0 elsewhere Liu et al. (2010). We define the input data training pattern set as follows:

$$\Phi = \begin{bmatrix} \mathbf{X} & \widehat{\mathbf{Y}} \end{bmatrix} \tag{4.9}$$

We used Φ to train the SOM's prototypes and assumed that the embedded class information supports inter-class separation and intra-class compactness throughout the learning. Once the training phase ends, the prototype's label is determined with a major voting criterion, using the class information attached in Φ as a reference. Figure 4.5 depicts an example of the clustering of the SOM's prototypes. After the class assignation finishes, it is necessary to remove the extended components from the prototypes so that their feature vectors agree with those of the voxels that will be relocated and have no embedded information.
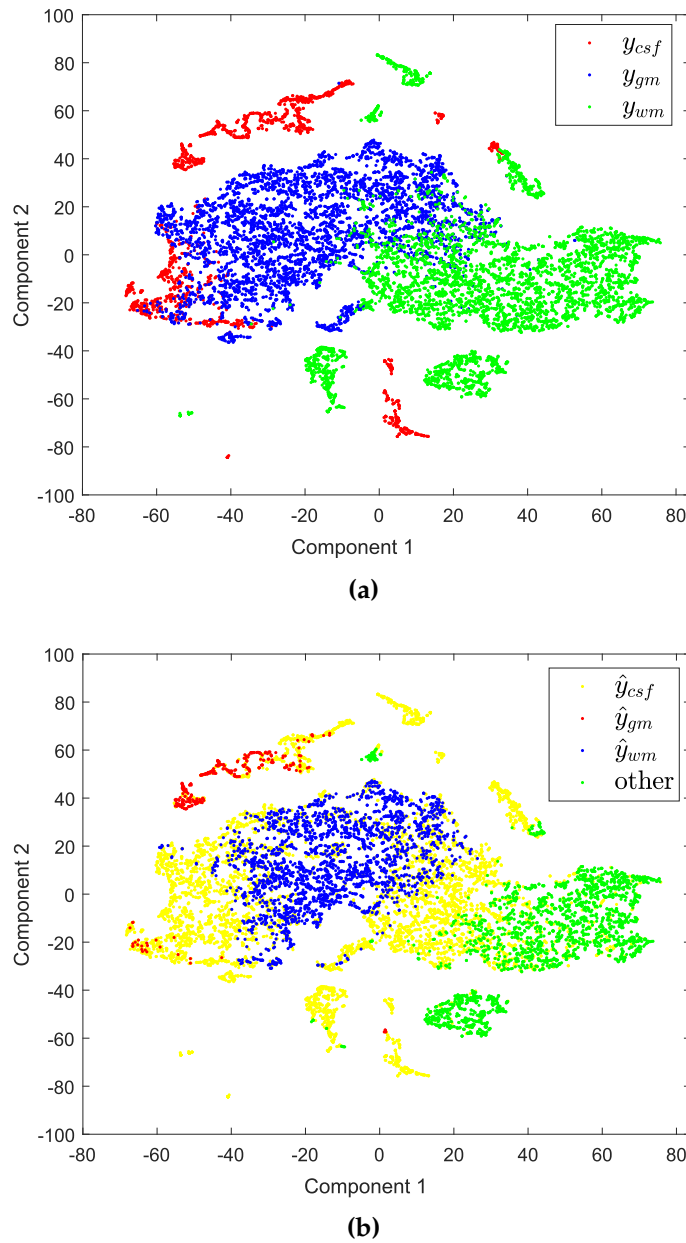
Figure 4.4: Example of the pseudo-label selection on the BrainWeb data set. a) data distribution using the ground truth. b) data distribution using the output from Gardens2.

4.3.3   The fuzzy membership calculation

According to the results from Eq. 4.8, we propose a novel membership computation to reallocate those voxels with a broad mixture level. The novel

(a)   **(b)**
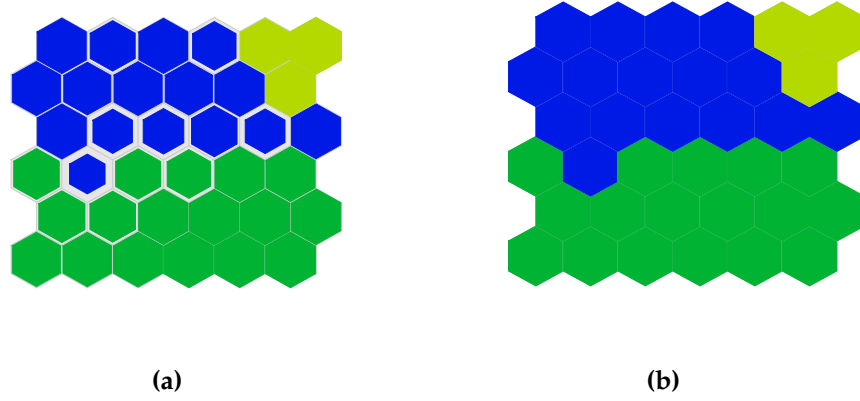
Figure 4.5: Similarity coloring graph for SOM's prototypes. (a) Hit map: the larger the
color area in each node, the more input patterns from the dominant class
hit it.(b) Clustered SOM nodes on the competitive layer.

membership computation is included in the SOM mapping phase and replaced
the traditional winner takes all approach. Instead of cluster the input patterns
directly with the BMUs, we use the membership summation from each proto-
type to every class to define the input pattern label. The rationale behind this
strategy is based on how the network prototypes learn. Once an input pattern
$\mathbf{x}_j$ meets its BMU_{j_x}, the neighborhood prototypes around the BMU_{j_x} receive
a stimulus to be more alike BMU_{j_x} to gather other input patterns similar to
$\mathbf{x}_j$. Analogously, we assumed that using a set of neighborhood prototypes
instead of a single prototype could improve the mapping phase. We intend to
take full advantage of the data distribution coded in the prototypes and their
ordering on the map. This clustering approach is useful on the SOM clusters'
border prototypes, where the uncertainty is larger than in the cluster's center.
Membership level of each prototype is computed as follows:

$$\mu_{jk} = \frac{\left(\|\mathbf{x}_j - \mathbf{w}_k\|^2 + \beta \sum_{l \in B_k} \sum_{m \in M_k} \mu_{lm}^q + \gamma \sum_{l \in N_j} \sum_{m \in M_k} \widehat{\Gamma}_{lm}^q\right)^{-\frac{1}{(q-1)}}}{\sum_{j=1}^{C} \left(\|\mathbf{x}_j - \mathbf{w}_k\|^2 + \beta + \sum_{l \in B_k} \sum_{m \in M_k} \mu_{lm}^q + \gamma \sum_{l \in N_j} \sum_{m \in M_k} \widehat{\Gamma}_{lm}^q\right)^{-\frac{1}{(q-1)}}} \quad (4.10)$$

where $\mu_{j,k}$ denotes the membership of the j-th voxel for a class k, x_j cor-
responds to the feature vector of the j-th voxel, $\mathbf{w}_k$ is the prototype of the
BMU for the j-th voxel. B_k represents the neighbor prototypes centered on

---

Input :$\mathbf{x}, \widehat{\Gamma}, k$

Output: brain tissue segmentation

1 Compute pseudo-labels $\widehat{y}$ with 4.8 ;

2 Build the embedded class information Φ with 4.9 ;

3 Train the model ;

4 Reallocate those voxels with a broad mixture level using 4.10;

Algorithm 2: PLA-SOM

---

the BMU for the j-th voxel. N_j is the voxels' neighborhood, in a 3D lattice, centered on the j-th voxel, $M_k = \{1, ..., C\}\backslash\{k\}, \{k \in C|C = \{csf, gm, wm\}\}$. The scalar parameters q, β, and γ control the degree of fuzziness, the amount of topological information from SOM and the amount of prior tissue information from $\widehat{\Gamma}$, respectively. The inclusion of prior tissue information corrects the possible errors made during the training phase, and provides the model with anatomical information in the form of spatial constraints. The values of $\mu_{j,k}$ are constrained to $[0 \quad 1]$ and satisfy $\sum_{k=1}^{C} \mu_{j,k} = 1$. Algorithm 2 summarizes the entire process.

Figure 4.6 depicts an example of the results from Eq. 4.10. Following the criterion winner takes all, an input pattern is mapped to $\mathbf{w}_{13}$. Hence, it will be labeled as WM, but the correct label of the input pattern is GM. From Figure 4.6, we can see that $\mathbf{w}_{13}$ is on the border of its cluster, and prototypes from other clusters surround it. When executing equation Eq. 4.10, it is observed that the neighboring prototypes around $\mathbf{w}_{13}$ have a membership degree to every cluster (blurred color), but their summation is greater for the GM cluster ($\mu = 0.66$). Therefore, using a maximum membership criterion, the input data will be labeled as GM.

The previous example shows two limitations in the mapping phase performed by the standard SOM model. The first limitation is related to uncertainty in the border prototypes of the SOM clusters. Regardless of the strategy used to cluster the prototypes, border elements are more prone to gather patterns from different classes. This phenomenon generates uncertainty in the clustering duty of such prototypes. Consequently, these elements are more prone to make mistakes in the mapping phase. The second limitation concerns the omission of topological information. Standard mapping in the SOM model is feature-based

---
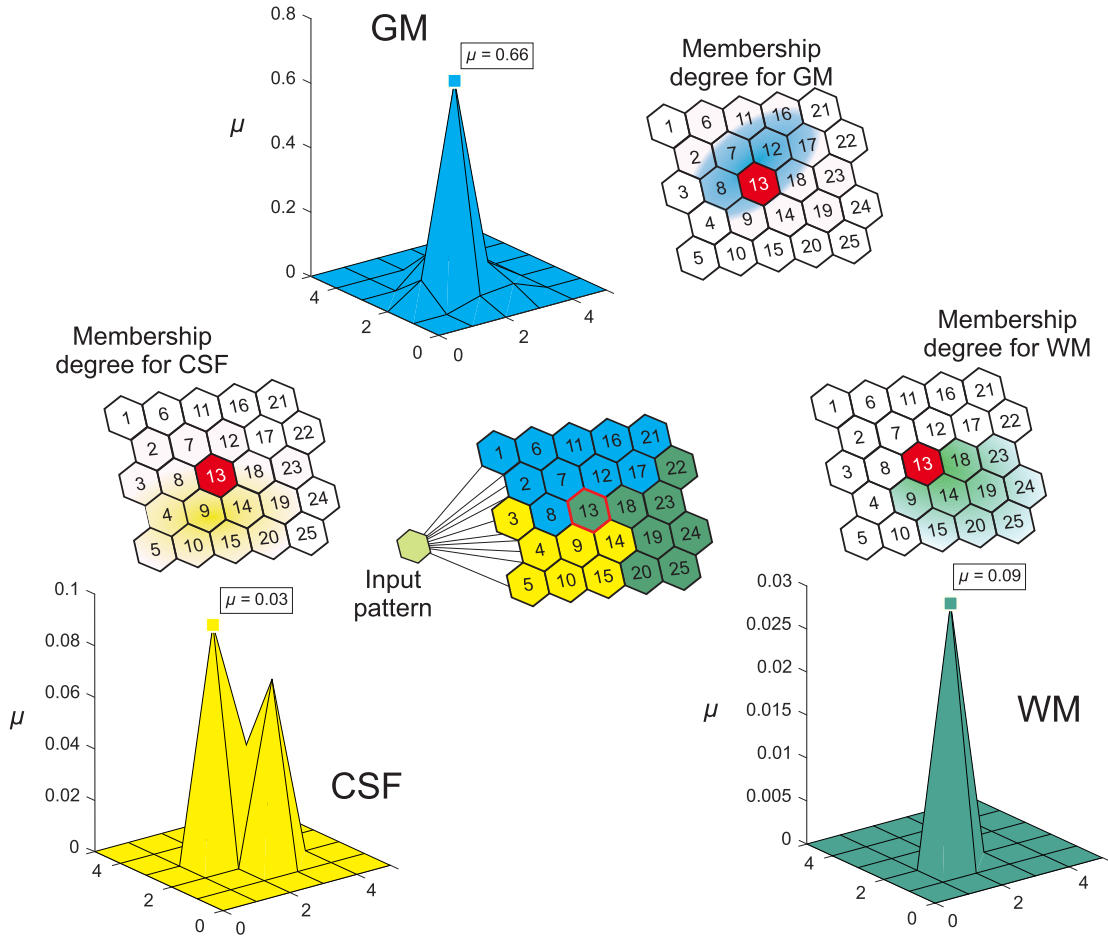
Figure 4.6: Graphical example of the label computation on PLA-SOM.

and ignores topological information learned by the model. Such knowledge can be useful to improve the mapping process and improve the SOM model's accuracy performance.

4.4 Summary

In this chapter, we introduced our two algorithms Gardens2 and PLA-SOM. Gardens2 adjusts the CSF, GM, and WM generic tissue maps to a specific target using a novel fuzzy function that combines two domains, feature and spatial. On the other hand, PLA-SOM presents a novel mapping procedure where combines prior knowledge (from Gardens2's output) and the data distribution learned by the prototypes and topological sorting from a SOM model.

EXPERIMENTAL RESULTS

This chapter describes the neuroimaging databases used in this research to validate Gardens2 and PLA-SOM algorithms. The validation covers the experimental setup, the experiments performed, the results, and the discussion.

5.1 MRI data

We assessed our method on two neuroimaging databases: BrainWeb Aubert-Broche et al. (2006) and Internet Brain Segmentation Repository (IBSR) IBSR (2007). The BrainWeb database provides synthetic brain MRI phantoms under well-characterized conditions of noise.

The IBSR dabase is divided into IBSR18 and IBSR20. This database allow us

Table 5.1: Description of the IBSR18 dataset.

IBSR18	
Number of subjects	18
Age	Average subject age was 38 years with an age range of $7-71$ years
Number of males/female	4 females, 14 males
Additional information	This dataset was created in 2003 and 2004. This dataset is a part of the Child and Adolescent NeuroDevelopment Initiative (CANDI) Kennedy et al. (2012)
Data acquisitions	
MRI system	1.5 T (Sonata, Siemens, Munich, Germany) and 1.5 T SIEMENS (Signa, General Electric, Boston, MA, USA)
MRI acquisition	

T1_w from IBSR18 present three different anisotropic voxel sizes ($0.84 \times 0.84 \times 1.5$ mm^3, $0.94 \times 0.94 \times 1.5$mm^3, and $1.0 \times 1.0 \times 1.5$mm^3). Scan parameters are unspecified. All the volumes are in Talairach orientation. Ground truth for this dataset is provided by the Center for Morphometric Analysis (CMA) at Massachusetts General Hospital Rohlfing (2011).

Table 5.2: Description of the IBSR20 dataset.

IBSR20	
Number of subjects	20
Age	Range of $20 - 38$ years
Number of males/female	Adults (10 females, 10 males)
Additional information	This dataset was created in 1997
Data acquisitions	
MRI system	1.5 T Siemens Magnetom MR System (Iselin, NJ) and 1.5 T General Electric Signa MR System (Milwaukee, WI)

MRI acquisition

Images are coronal three-dimensional $T1_w$ spoiled gradient echo MRI scans, that were obtained on two different imaging systems. 10 FLASH scans performed on a with the following parameters: TR = 40 ms, TE = 8 ms, $\alpha = 50$, FOV = 30 cm, and anisotropic voxel size $1.0 \times 1.0 \times 3.1$ mm^3. 10 3D-CAPRY scans with the following parameters: TR = 50 ms, TE = 9 ms, $\alpha = 50°$, FOV = 24 cm, and anisotropic voxel size $1.0 \times 1.0 \times 3.1$ mm^3. All the volumes are in Talairach orientation. The MR brain images in this dataset and their manual segmentations were provided by the CMA at Massachusetts General Hospital IBSR (2007).

Table 5.3: Description of the BrainWeb dataset.

BrainWeb	
Number of subjects	20
Age	Average subject age was 29.6 years with an age range of $24 - 37$ years
Number of males/female	Adults (10 females, 10 males)
Additional information	The Ethics Committee of the Montreal Neurological institute and Hospital approved the study, and written informed consent was obtained from all participants
Data acquisitions	
MRI system	1.5 T Siemens Sonata Vision clinical scanner (Siemens Medical Systems, Erlangen, Germany)

MRI acquisition

The protocol comprises conventional whole-head high-resolution $T1_w$, $T2_w$, PD_w scans, and a magnetic . resonance angiography (MRA) scan. The $T1_w$ scan was acquired using a 3D spoiled gradient echo (GRE) sequence (TR = 22 ms, TE = 9.2 ms, and $\alpha = 30$), providing whole head coverage with 1.0 mm isotropic voxel size. The PD_w and $T2_w$ images were acquired within the same turbo spin-echo (TSE) protocol (TR = 3.3 s, TE = 15/104 ms), providing whole head coverage in 80 2.0 mm slices, with 1.0×1.0 mm in-plane resolution. The MRA was acquired using a 3D phase-contrast spoiled gradient-echo angiography scan with 176 0.9 mm slices (TR = 71 ms, TE = 8.2 ms, and $\alpha = 15°$), with 0.47×0.47 mm in-plane resolution. For each subject, each $T1_w$, $T2_w$, and PD_w scans scan was repeated four times in order to increase the signal-to-noise ratio (SNR) by averaging, after registration. All the volumes are in Talairach orientation. Ground truth for this dataset is provided by BrainWeb Aubert-Broche et al. (2006)

Data modifications

We corrupted the $T1_w$ brain phantoms by adding 7% Rician of noise level.

to compare our method's performance with other methods that addressed the same problem on real data. Tables 5.1, 5.2, and 5.3 describe each dataset.

5.2   Experiment 1: Performance of Gardens2

This experiment aims to assess the performance of Gardens2 for adjusting generic atlas tissue templates (CSF, GM, and WM) to a specific subject. The experimental hypothesis states that the correspondence between the atlas and the target can improve if the similarity measure includes spatial constraints and features that describe in more detail the gray-intensity relationships in the image (3D Haralick texture features). For the parameters configuration, we set $\omega = 3$ and 16 gray levels for the co-occurrence matrix, $q = 2$ and $\gamma = 0.01$ for Gardens2. These parameters were set following experimental results.

We assessed Garden2 on brain phantoms from BrainWeb, because they are realistic images created with an MRI simulator on human brain ground truths. We compared the performance of Gardens2 with the BRAINS algorithm, which is a tool provided by the 3D Slicer software platform Fedorov et al. (2012). In this research, the latest 3D Slicer version was used for co-registering the atlas tissue templates. We implemented the BRAINS algorithm with default parameters recommended by their authors. We took a sample of 50 consecutive MRI brain scans, twenty-five upwards and downwards from the middle section of the brain in the axial plane, from each subject. We used the RMSE to assess the performance of each method. Next, we applied the Wilcoxon signed-rank test Wilcoxon (1992), with Bonferroni correction to find significant differences between the performance of both methods with $\alpha = 0.05$. We complement the Wilcoxon test information with the effect size computed from Cohen's D Cohen and Hillsdale (1988) and confidence interval (CI). We used the scale suggested by Cohen to measure the effect size since it provides a straightforward interpretation, and it is a conventional frame of reference. For Cohn's D a value of 0.1 is considered *small*, 0.3 is a *medium* effect size and 0.5 a *large* effect size. In this study, we used a confidence level of 95%.

5.2.1   Results on BrainWeb

Table 5.4 shows the statistical results for Gardens2 and BRAINS algorithm evaluated over brain MRI phantoms. Gardens2 achieved the best RMSE results for CSF (0.042 ± 0.006), GM (0.068 ± 0.009), and WM (0.057 ± 0.008). On the Overall performance, Gardens2 reported the best result (0.056 ± 0.006), while BRAINS achieved the second-best result (0.132 ± 0.016). The inferential analysis revealed a global difference between the results of Gardens2 and BRAINS because **Wilcoxon signed-rank:** $4.42 \times 10^{-5} < 0.05$ and **Cohen's D:** 0.87 (large), 95% CI 0.87 to 0.88.

We were unable to compare the Gardens2 performance with other co-registration methods such as those included in FAST, US, and PVC since the results of such procedures are not available as an output. These methods consider their segmentation result as the system's main output and do not provide intermediate results as the co-registration process. Additionally, in Appendix A, Section A.1, we present the segmentation results of Gardens2. Figure 5.1 depicts an example of the output from Gardens2 given a set of initial prior tissue models from a generic atlas template for CSF, GM, and WM.

Table 5.4: RMSE results for CSF, GM, and WM volumes on BrainWeb dataset.

Subject	CSF		GM		WM		Overall	
	BRAINS	Gardens2	BRAINS	Gardens2	BRAINS	Gardens2	BRAINS	Gardens2
subject 1	0.064	0.036	0.161	0.057	0.136	0.046	0.120	0.047
subject 2	0.069	0.036	0.150	0.055	0.132	0.044	0.117	0.045
subject 3	0.061	0.038	0.155	0.081	0.124	0.075	0.113	0.065
subject 4	0.063	0.038	0.172	0.073	0.152	0.070	0.129	0.060
subject 5	0.070	0.036	0.163	0.060	0.156	0.049	0.129	0.048
subject 6	0.075	0.038	0.169	0.057	0.147	0.050	0.130	0.048
subject 7	0.086	0.038	0.153	0.069	0.121	0.058	0.120	0.055
subject 8	0.071	0.037	0.177	0.064	0.182	0.057	0.144	0.053
subject 9	0.073	0.034	0.170	0.058	0.156	0.055	0.133	0.049
subject 10	0.083	0.040	0.162	0.057	0.157	0.049	0.134	0.049
subject 11	0.077	0.048	0.172	0.072	0.160	0.062	0.136	0.060
subject 12	0.061	0.052	0.164	0.075	0.144	0.054	0.123	0.060
subject 13	0.087	0.050	0.190	0.083	0.176	0.063	0.151	0.066
subject 14	0.118	0.051	0.233	0.070	0.196	0.055	0.182	0.059
subject 15	0.062	0.043	0.155	0.060	0.159	0.050	0.126	0.051
subject 16	0.041	0.035	0.169	0.071	0.154	0.057	0.121	0.054
subject 17	0.062	0.048	0.180	0.073	0.166	0.059	0.136	0.060
subject 18	0.099	0.043	0.169	0.079	0.154	0.057	0.141	0.060
subject 19	0.087	0.044	0.200	0.079	0.178	0.059	0.155	0.061
subject 20	0.063	0.052	0.177	0.075	0.164	0.059	0.135	0.062
	0.074 ± 0.017	0.042 ± 0.006	0.172 ± 0.019	0.068 ± 0.009	0.156 ± 0.019	0.057 ± 0.08	0.134 ± 0.016	0.056 ± 0.006

| | CSF | GM | WM |
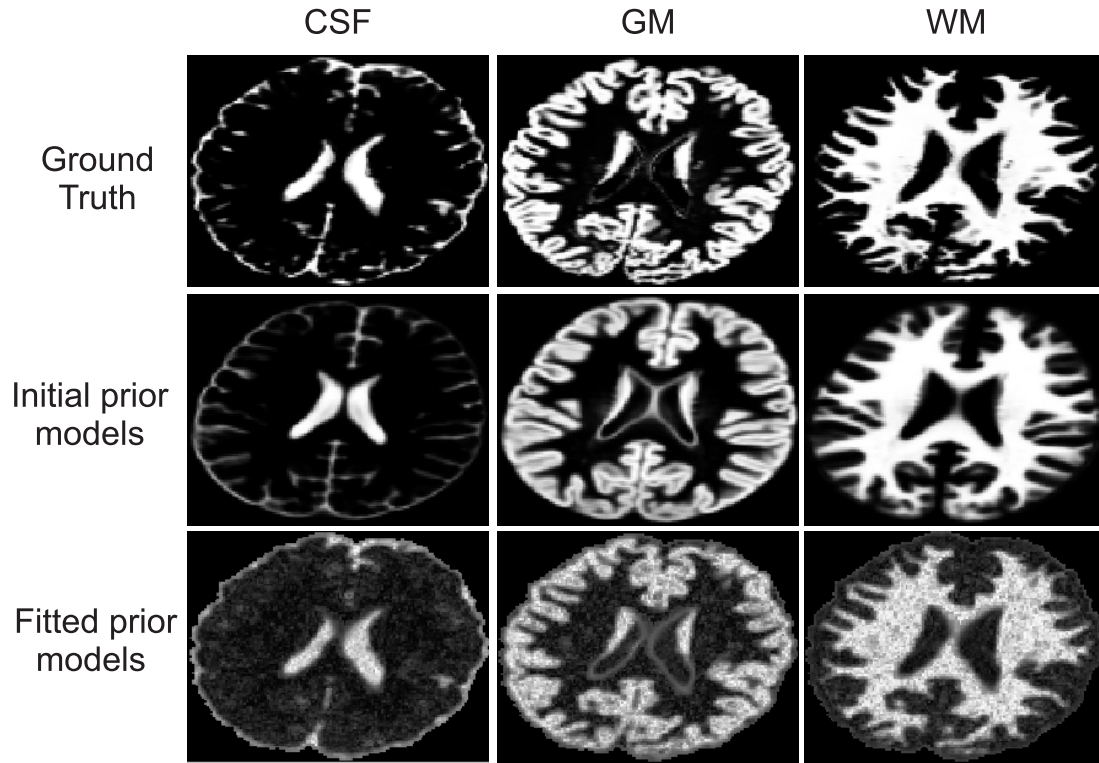
Figure 5.1: Comparison of the initial atlas co-registration (initial prior tissue models) and the output from Gardens2.

5.3 Experiment 2: Performance comparison of PLA-SOM

This experiment aims to assess the performance of the PLA-SOM algorithm for segmenting brain tissues (CSF, GM, and WM). The experimental hypothesis states that a method's segmentation performance can improve if the model learns the features of voxels that have a major relationship toward a particular class, for then cluster those voxels with high uncertainty. PLA-SOM uses the output from Gardens2 as a prior knowledge as well as input to compute the pseudo-labels. For the parameters configuration, we set $\tau_p = 0.8$, $q = 2$, $\beta = 0.07$, and $\gamma = 0.05$, this parameters were set following experimental results.

We assessed the PLA-SOM algorithm's segmentation performance on three neuroimaging datasets: BrainWeb, IBSR18, and IBSR20. These datasets present different resolutions and acquisition artifacts (noise, bias field, and TFE) and

provide a broad diversity to test our method's reliability. We compared the segmentation performance of Gardens2-PLA-SOM with the results from five well-established brain tissue segmentation methods, namely FAST, US, PVC, and two extensions of the FCM algorithm known as the Fuzzy and Noise Tolerant Adaptive Segmentation Method (FANTASM) Pham (2001a) and the FCM_S1 algorithm Chen and Zhang (2004). We implemented all the above methods using default parameters recommended by their authors. We also compared the results of other current methods that have been tested before on the selected datasets for this study.

We used the same images that were described in experiment 1. For the part of the descriptive statistics, we report different measures such as minimum (Min), maximum (Max), median (Median), quartile 1 (Q1), quartile 3 (Q3), interquartile range, (IQR), mean (Mean), and standard deviation (SD). These statistics were computed from the DSC mean result of each method on each subject. For the inferential analysis, we first applied the Friedman test Friedman (1940) to detect significant differences, with $\alpha = 0.05$, among the segmentation results provided by multiple methods. We measured the Friedman test's effect size with Kendall's W Kendall and Gibbons (1990) and confidence interval (CI). Finally, we applied the Wilcoxon signed-rank test Wilcoxon (1992) with Bonferroni correction as the post-hoc procedure. The Bonferroni correction was used to adjust p-value to reduce the chance of observing a rare event (Type I error) under multiple comparisons. We used the scale suggested by Cohen, described in experiment 1, to measure the effect size for Kendall's W a value.

5.3.1   Results on BrainWeb

Table 5.5 shows the results obtained by all the compared methods in this study for each tissue class. The best results are in bold. PLA-SOM achieved the best performance (0.77 ± 0.04) on CSF class. The second-best result for the CSF class was reported by FANTASM (0.75 ± 0.09). On the GM class, PLA-SOM reported the best result (0.83 ± 0.02), followed by US (0.78 ± 0.03). However, the best result for the WM class was achieved by US (0.85 ± 0.02), followed by PLA-SOM (0.84 ± 0.01). Table 5.6 contains the descriptive statistics, PLA-SOM reported the best results for the mean (0.82 ± 0.04), median (0.82) and IQR (0.01). Box plot depicted on Figure 5.3 complement descriptive statistic

analysis. From Table 5.7, Friedman test revealed significant differences among compared brain tissue segmentation methods (p-value<0.05) and large effect size ($r > 0.5$). Wilcoxon's test evidenced the statistical difference ($p < 0.5$) between PLA-SOM against the other methods. From Tables 5.5, 5.6, and 5.7, we conclude that PLA-SOM presents the best results compared with the other methods for segmenting brain tissues on BrainWeb dataset with a Rician noise

Table 5.5: Results per class obtained on the BrainWeb for the DSC measure (mean ± standard deviation).

Method	CSF	GM	WM
FANTASM Pham (2001a)	0.75 ± 0.09	0.74 ± 0.02	0.80 ± 0.02
FAST Zhang et al. (2001)	0.71 ± 0.13	0.75 ± 0.07	0.81 ± 0.05
FCM_S1 Chen and Zhang (2004)	0.63 ± 0.20	0.71 ± 0.11	0.75 ± 0.10
PLA-SOM	$\mathbf{0.77 \pm 0.04}$	$\mathbf{0.83 \pm 0.02}$	0.84 ± 0.01
PVC Shattuck et al. (2001)	0.48 ± 0.29	0.52 ± 0.15	0.73 ± 0.05
US Ashburner and Friston (2005)	0.68 ± 0.09	0.78 ± 0.03	$\mathbf{0.85 \pm 0.02}$
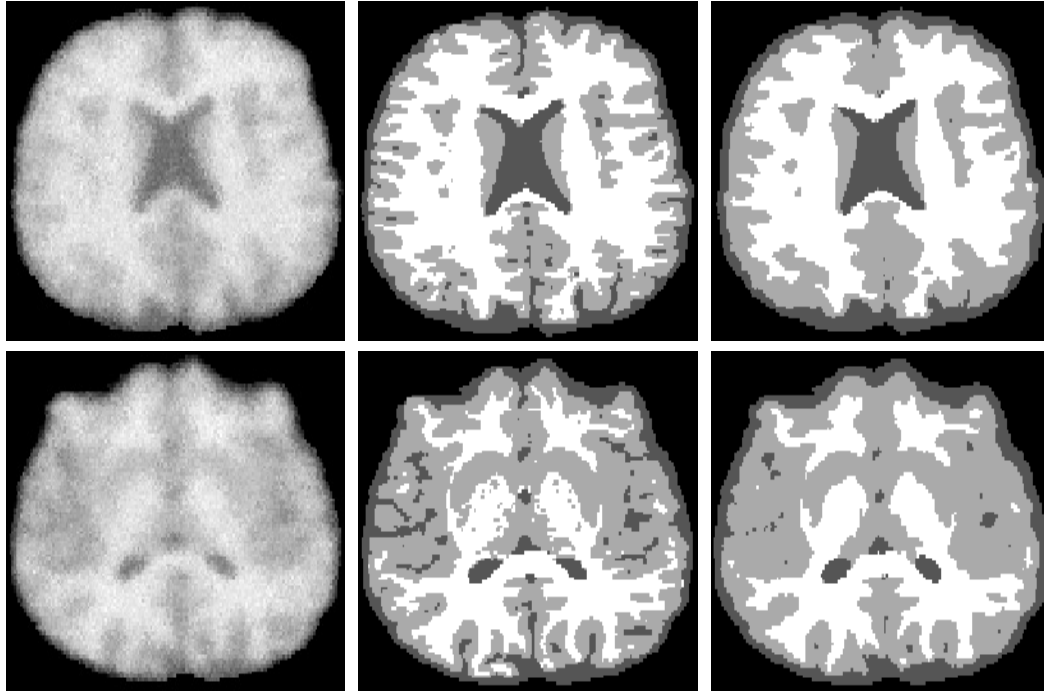
Figure 5.2: Segmentation results for volume 20 in dataset BrainWeb. From left to right: $T1_w$ scan, ground truth and segmentation obtained with PLA-SOM.

Table 5.6: Descriptive statistics on BrainWeb.

Methods	Min	Max	Median	Q1	Q3	IQR	Mean
FANTASM Pham (2001a)	0.60	0.79	0.77	0.76	0.78	0.02	0.76 ± 0.04
FAST Zhang et al. (2001)	0.47	0.82	0.79	0.77	0.80	0.03	0.76 ± 0.08
FCM_S1 Chen and Zhang (2004)	0.44	0.82	0.79	0.57	0.81	0.24	0.70 ± 0.13
PLA-SOM	0.66	0.85	0.82	0.82	0.83	0.01	0.82 ± 0.04
PVC Shattuck et al. (2001)	0.38	0.77	0.56	0.44	0.75	0.31	0.58 ± 0.16
US Ashburner and Friston (2005)	0.63	0.81	0.77	0.76	0.79	0.03	0.77 ± 0.04
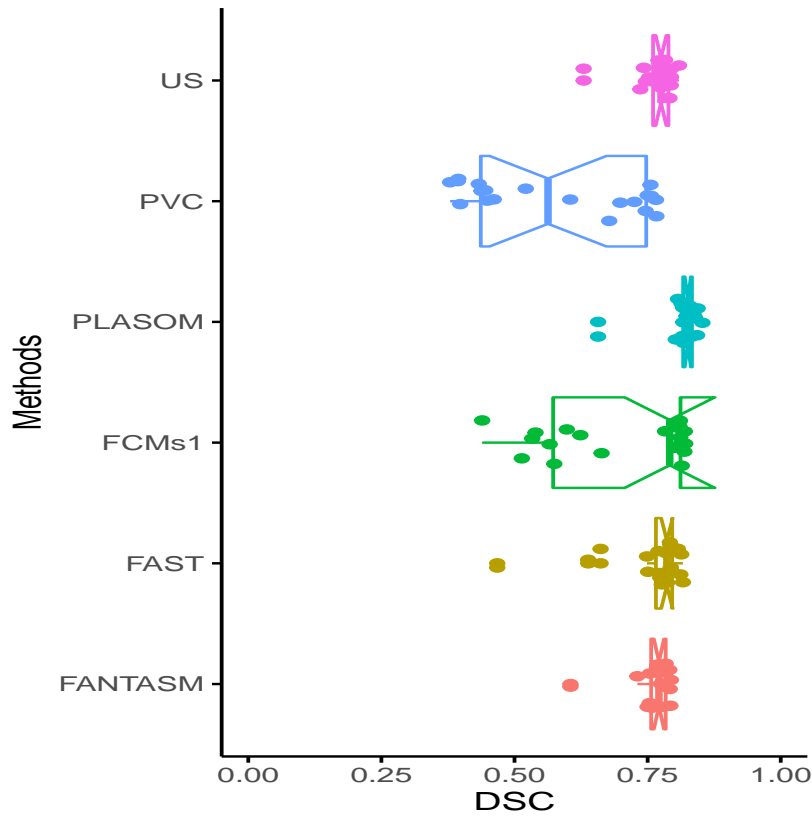
Figure 5.3: Box plot for the DSC from the methods on BrainWeb.

of 7% and no bias field inhomogeneity condition.

5.3.2　Results using IBSR18 and IBSR20

The IBSR18 dataset provides the brain MRI scans already skull-stripped and bias field corrected. Therefore, image intensity normalization was the only

Table 5.7: Inferential statistical analysis for all compared methods on BrainWeb.

Friedman test: $\chi^2 = 53.464$, df = 5, $p\text{-val} = 2e - 10$	
Kendall's W: 0.53 (large), 95% CI 0.49 to 0.65	

Post-hoc Wilcoxon signed-rank:	
Comparison	p**-value**
PLA-SOM vs FANTASM	9.5×10^{-5}
PLA-SOM vs FAST	1.9×10^{-6}
PLA-SOM vs FCM_S1	2.7×10^5
PLA-SOM vs PVC	1.9×10^{-6}
PLA-SOM vs US	8.2×10^{-5}

preprocessing procedure applied to the IBSR18 dataset. The IBSR20 dataset volumes were skull-stripped, but they have a severe noise and bias field inhomogeneity condition. Hence, we applied the N3 algorithm Sled et al. (1998) and the Non-local Filtering algorithm to reduce the artifacts (noise and bias field) in the volumes. We performed the rest of the procedures of the segmentation pipeline of PLA-SOM without any further modifications. We used the original manual segmentation on both IBSR datasets, where the sulcus CSF is part of the ground truth Valverde et al. (2015).

Table 5.8 shows the results obtained by all the compared methods in this study for each tissue class on the IBSR datasets; * indicates that the results were taken from the corresponding publications; NA indicates that the performance of the algorithm for this tissue was not reported. The best results are in bold. We first describe the results of PLA-SOM on the IBSR18 dataset. PLA-SOM reported the best result (0.73 ± 0.04) for the CSF class, followed by the PVC (0.58 ± 0.07). On the GM class, PLA-SOM reported the best result (0.85 ± 0.03), while the US achieved the second-best result (0.82 ± 0.03). FAST achieved the best result (0.89 ± 0.02) on the WM class. PLA-SOM reported one of the lowest results on the WM class (0.79 ± 0.04) but with a competitive agreement between manual and automatic segmentation Cabezas et al. (2011).

The IBSR20 volumes present low resolution and contrast and severe artifacts (noise and bias field), compared with BrainWeb and IBSR18 datasets.

Table 5.8: Results per class obtained on IBSR18 and IBSR20 datasets for the DSC measure (mean ± standard deviation).
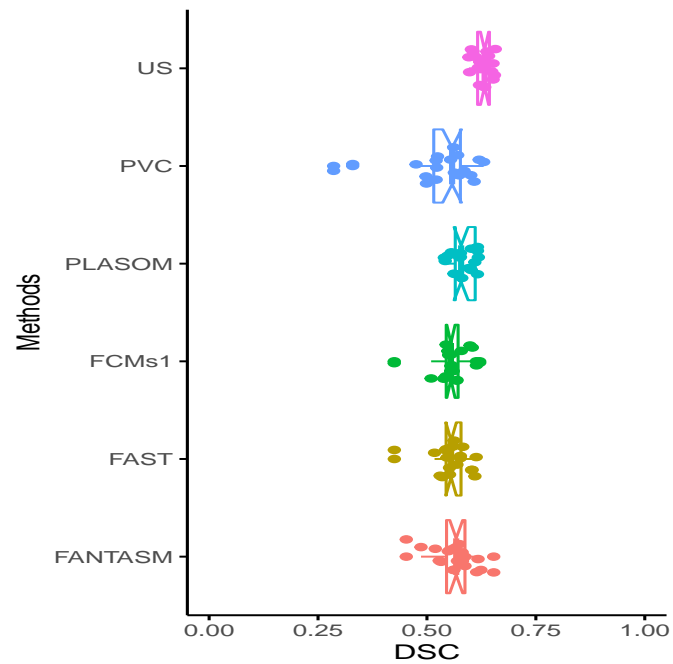
IBSR18			
Method	CSF	GM	WM
FAST Zhang et al. (2001)	0.57 ± 0.06	0.80 ± 0.05	$\mathbf{0.89 \pm 0.02}$
US Ashburner and Friston (2005)	0.48 ± 0.09	0.82 ± 0.03	0.86 ± 0.01
FANTASM Pham (2001a)	0.45 ± 0.11	0.72 ± 0.07	0.88 ± 0.03
PVC Shattuck et al. (2001)	0.58 ± 0.07	0.75 ± 0.08	0.85 ± 0.07
FCM_S1 Chen and Zhang (2004)	0.44 ± 0.12	0.70 ± 0.08	0.86 ± 0.03
EGS-SOM Ortiz et al. (2013a)*	0.36 ± 0.08	0.82 ± 0.04	0.82 ± 0.04
HFS-SOM Ortiz et al. (2013a)*	0.18 ± 0.05	0.75 ± 0.15	0.75 ± 0.10
FCCNN Agnello et al. (2016)*	0.58 ± 0.11	0.81 ± 0.03	0.84 ± 0.05
T-HMM Shenoy et al. (2016)*	NA	0.78 ± 0.02	0.76 ± 0.02
PLA-SOM	$\mathbf{0.73 \pm 0.04}$	$\mathbf{0.85 \pm 0.03}$	0.79 ± 0.04

IBSR20			
Method	CSF	GM	WM
FAST Zhang et al. (2001)	0.18 ± 0.05	0.70 ± 0.04	0.80 ± 0.06
US Ashburner and Friston (2005)	$\mathbf{0.24 \pm 0.06}$	0.81 ± 0.03	0.83 ± 0.03
FANTASM Pham (2001a)	0.20 ± 0.07	0.71 ± 0.06	0.79 ± 0.05
PVC Shattuck et al. (2001)	0.17 ± 0.06	0.68 ± 0.07	0.74 ± 0.19
FCM_S1 Chen and Zhang (2004)	0.20 ± 0.07	0.70 ± 0.04	0.77 ± 0.03
FAMS Mayer and Greenspan (2009)*	NA	0.82 ± 0.03	0.82 ± 0.04
EGS-SOM Ortiz et al. (2013a)*	NA	$\mathbf{0.84 \pm 0.04}$	$\mathbf{0.86 \pm 0.05}$
HFS-SOM Ortiz et al. (2013a)*	NA	0.75 ± 0.09	0.75 ± 0.08
SCSOTM Zhang and Jiang (2014)*	NA	0.72 ± 0.03	0.65 ± 0.06
PLA-SOM	0.23 ± 0.05	0.78 ± 0.02	0.74 ± 0.04

Those issues negatively affect the prior tissue fitting procedure; Watershed transformation produces many errors when defining Watershed lines. The tissue interface GM/WM is the most affected by this problem; this output produces an inappropriate estimation on the pseudo labels for PLA-SOM. The outlined problem shows up because we assumed that prior information from an atlas would be enough to overcome the artifacts present in a brain MRI scan. Therefore, we omitted spatial information (membership level of neighboring voxels) from the image in the Gardens2 procedure (See Eq. 4.2). However, from Table 5.8 we can see that PLA-SOM reported a competitive result for CSF (0.23 ± 0.05), GM (0.78 ± 0.02), and WM (0.74 ± 0.04), compared with the result
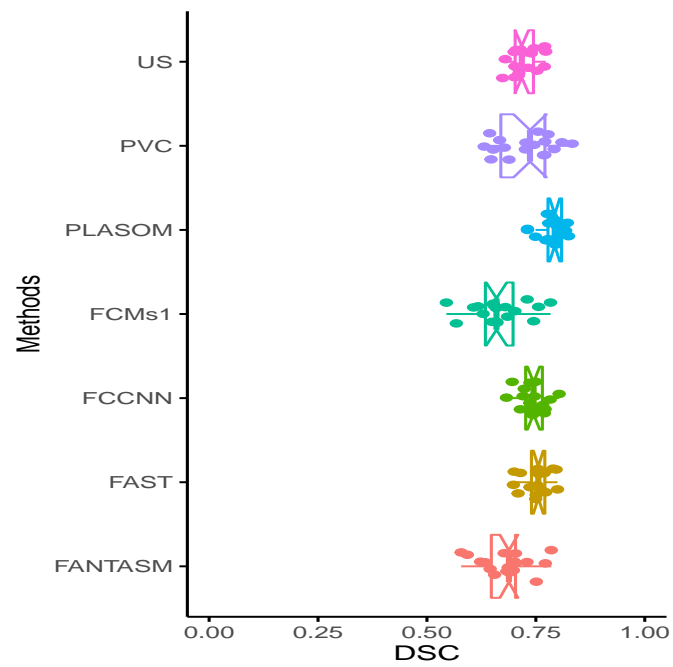
(a)

(b)

Figure 5.4: Box plot for the DSC from the methods applied on the IBSR18 and IBSR20 datasets.

from other methods.

Table 5.9 contains the descriptive statistics obtained by all the compared methods in this study for each tissue class on the IBSR dataset. We first describe the results on the IBSR18 dataset. PLA-SOM reported the best results for the mean (0.79 ± 0.02), median (0.77), and narrowest dispersion IQR (0.03). Box plot depicted on Figure 5.4(a) complement descriptive statistic analysis. From Tables 5.8 and 5.9, note that the performance of methods without atlas information (FCNN, EGS-SOM, and FANTASM) reach the performance of methods with atlas information (FAST, US, and PVC). This particular performance is the result of the high resolution and image quality on IBSR18 scans. Furthermore, scans got through to a correction routine; gray-intensity correction and missing CSF labels were added to previously unlabeled regions Rohlfing (2011). As a result, the brain scans do no present severe acquisition artifacts (noise and bias field) that can bias the methods' accuracy. For the IBSR20 dataset, from Table 5.9, we can see that PLA-SOM achieved the second-best results for the mean (0.58 ± 0.03), median (0.58), and IQR (0.05). Box plot depicted on Figure 5.4(b)

Table 5.9: Descriptive statistics on IBSR18 and IBSR20.

IBSR18							
Methods	Min	Max	Median	Q1	Q3	IQR	Mean
FANTASM Pham (2001a)	0.58	0.78	0.69	0.65	0.70	0.06	0.68 ± 0.06
FAST Zhang et al. (2001)	0.70	0.80	0.76	0.74	0.77	0.03	0.75 ± 0.03
FCNN Agnello et al. (2016)	0.68	0.80	0.74	0.73	0.76	0.04	0.74 ± 0.03
FCM_S1 Chen and Zhang (2004)	0.54	0.78	0.66	0.63	0.70	0.06	0.67 ± 0.06
PLA-SOM	0.73	0.82	0.79	0.78	0.81	0.03	0.79 ± 0.02
PVC Shattuck et al. (2001)	0.63	0.83	0.74	0.67	0.77	0.10	0.73 ± 0.06
US Ashburner and Friston (2005)	0.67	0.77	0.72	0.70	0.74	0.04	0.72 ± 0.03

IBSR20							
Methods	Min	Max	Median	Q1	Q3	IQR	Mean
FANTASM Pham (2001a)	0.45	0.65	0.57	0.54	0.59	0.04	0.56 ± 0.05
FAST Zhang et al. (2001)	0.42	0.61	0.55	0.54	0.58	0.03	0.56 ± 0.03
FCM_S1 Chen and Zhang (2004)	0.42	0.62	0.55	0.54	0.57	0.03	0.56 ± 0.04
PLA-SOM	0.54	0.62	0.58	0.56	0.61	0.05	0.58 ± 0.03
PVC Shattuck et al. (2001)	0.29	0.63	0.56	0.52	0.58	0.06	0.53 ± 0.09
US Ashburner and Friston (2005)	0.60	0.66	0.63	0.62	0.64	0.03	0.63 ± 0.02

Table 5.10: Inferential statistical analysis for all compared methods applied on IBSR18 and IBSR20 dataset.

IBSR18	
Friedman test: $\chi^2 = 56.5$, df = 6, p-val $= 2.26e - 10$	
Kendall's W: 0.52 (large), 95% CI 0.39 to 0.71	
Post-hoc Wilcoxon signed-rank:	
Comparison	**p-value**
PLA-SOM vs FAST	5.34×10^{-4}
PLA-SOM vs US	2.29×10^{-5}
PLA-SOM vs FANTASM	5.34×10^{-5}
PLA-SOM vs PVC	5.0×10^{-3}
PLA-SOM vs FCM_S1	1.15×10^{-5}
PLA-SOM vs FCCNN	1.91×10^{-4}
IBSR20	
Friedman test: $\chi^2 = 59.1$, df = 5, p-val $= 1.88e - 11$	
Kendall's W: 0.59 (large), 95% CI 0.52 to 0.73	
PLA-SOM vs FANTASM	0.4
PLA-SOM vs FAST	0.01
PLA-SOM vs FCM_S1	2.0×10^{-3}
PLA-SOM vs PVC	7.1×10^{-4}
PLA-SOM vs US	1.91×10^{-6}

complement descriptive statistic analysis.

Table 5.10 summarizes the inferential statistical results. We found significant differences among the compared algorithms' segmentation performance on the IBSR18; Friedman test showed p-value< 0.05 and large effect size (> 0.5). We found significant differences between PLA-SOM's segmentation performance and the other methods; the Wilcoxon test showed p-value< 0.05. There is a particular situation for the comparison between PLA-SOM and PVC. This comparison revealed a p-value$> \alpha$. We could not apply the complete statistical analysis (descriptive and inferential) on HFS-SOM, EGS-SOM, neither T-HMM methods because they only reported their mean results on each tissue class. We

found significant differences among the compared algorithms' segmentation performance on the IBSR20; Friedman test showed p-value< 0.05 and large effect size (> 0.5). In the pairwise comparison, we found that there is a significant difference between PLA-SOM and FAST, FCM_S1, PVC, and US (p-value$\leqslant 0.05$), but not for PLA-SOM and FANTASM (p-value> 0.05). We could not apply the complete statistical analysis (descriptive and inferential) on FAMS, HFS-SOM, EGS-SOM, neither SCSOTM methods because they only reported their mean results on each tissue class.

5.3.3 Discussion about the segmentation performance of Gardens2 and PLA-SOM

From the results, we can see that Gardens2 achieved a better tissue fitting on CSF, GM, and WM than a baseline co-registration method like BRAINS. These results provide evidence that the correspondence between the atlas and the target can improve if the similarity measure includes spatial constraints and features that describe in more detail the gray-intensity relationships in the image. Gardens2 introduced a novel split split-and-merge approach to circumvent relevant mismatches between feature and spatial features. In this manner, Gardens2 gets the benefit from the prior tissue model and bounds the cumulative error effect generated by the mismatch with irregular brain regions such as cortical sulci and gyri and intra-sulcus CSF. This approach is useful for decreasing the class bias during the pseudo-labels calculation.

On the other hand, we observed that PLA-SOM reported the best results on two (BrainWeb and IBR18) of the three datasets used in this research. These results provide evidence that a method's segmentation performance can improve if the model learns the features of voxels that have a major relationship toward a particular class for then cluster those voxels with high uncertainty. To the best of our knowledge, the novel mapping introduced by PLA-SOM is one of the few that incorporates the topological sorting learned by the model into the clustering. This strategy can reduce the uncertainty in the clustering phase and improve the SOM model's performance.

PLA-SOM addresses the segmentation of three major brain tissue (CSF, GM, and WM). It has a remarkable performance for segmenting CSF and GM
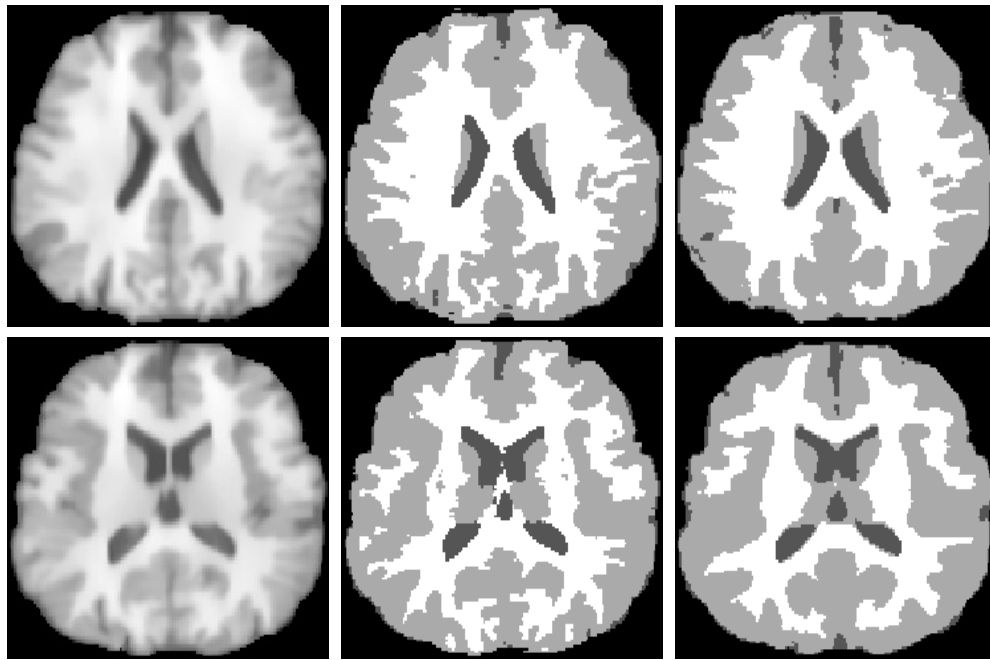
Figure 5.5: Segmentation results for the IBSR18, volume 18. From left to right: T1$_w$ scan, ground truth and segmentation obtained whit PLA-SOM.
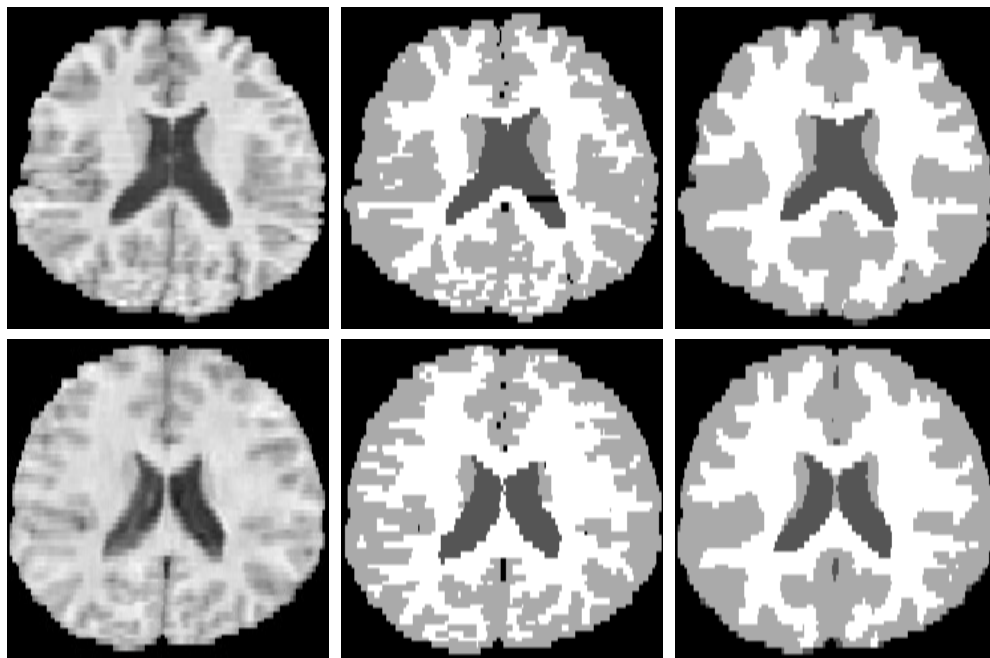

Figure 5.6: Segmentation results for the IBSR20, volume 19. From left to right: T1w scan, ground truth and segmentation obtained with PLA-SOM.

tissues, but it struggles to segment WM tissue. We can see from the results that none of the unsupervised methods studied obtained the best result for the three classes on the same database. This outcome relies on the fact that the datasets contain only $T1_w$ sequences. The different MRI sequences have some distinct signal intensity that causes structural distinctions to appear among the tissues. For example, a tissue can be visible clearly in $T2_w$ and embodied too little in $T1_w$, or, in turn, a tissue can be visible in $T1_w$ and is embodied too little in $T2_w$. These distinct tissues are no ghosts and maybe lesions, healthy tissues, or tissue interfaces. Therefore, the methods that address brain tissue segmentation should consider the multi-channel approach, a limitation that PLA-SOM has.

For clinical applications, studies suggest that a segmentation method should report an accuracy rate above 70% Cabezas et al. (2011); Porz et al. (2014). PLA-SOM meets this requirement for BrainWeb and IBSR18 datasets, but not for the IBSR20 dataset. PLA-SOM can support, as a secondary tool, the assessment of pathologies related to brain volume reduction. For example, progressive brain atrophy, characterized by a decrease in brain volume. Pathologies, such as brain tumors El-Dahshan et al. (2014); Subudhi et al. (2018), Multiple Sclerosis Shiee et al. (2010), or Alzheimer's disease Prasad et al. (2015), are out of the scope of PLA-SOM because the analysis of such neurological disorders demands MRI modalities additional to $T1_w$. Figures 5.2, 5.5, and 5.6 depict examples of brain tissue segmentation results obtained by PLA-SOM.

CONCLUSIONS

Brain tissue segmentation methods address the parcellation of brain tissues depicted on MRI scans, an ill-posed problem. Thus, there is a broad diversity of methods trying to provide reliable approximations for the brain tissue segmentation problem. The state-of-the-art methods assume that spatial information is the key to overcome the issues around the segmentation problem. However, spatial information management is not a trivial task, and different approaches lead to different outcomes.

This thesis presented a novel procedure to employ the available spatial information to design a robust and accurate brain tissue segmentation method based on a dichotomy of brain tissue segmentation methods. We built an algorithm named *Gardens2*, which adjusts the generic prior tissue models from an atlas to a specific target using a novel fuzzy function. This function estimates the membership degree for each voxel on every prior tissue model. Then, we proposed a new algorithm, called *Pseudo-Label Assisted Self-Organized Map (PLA-SOM)* to addresses the brain tissue segmentation problem. PLA-SOM improves the inter-class separation and an intra-class compactness in the training phase using pseudo-labels computed from Gardens2. Furthermore, PLA-SOM brings a novel mapping phase where the membership summation from each prototype associated with the input pattern defines the label. This new mapping approach is practical for dealing with uncertainty in border prototypes.

6.1   Research question

How to accurately adjust a model, which has been estimated by an atlas co-registration process, to a specific target?

We found that the correspondence between the atlas and the target subject improves if the feature and spatial spaces are combined. The split-and-merge strategy is useful for identifying thin brain structures and voxels affected by the TFE. Gardens2 fits better CSF and GM prior maps than WM. One of the main reasons for this result is because the overlapping between GM-WM is larger than in CSF-GM. However, this overlap only applies to $T1_w$ sequences. We consider that the correspondence can improve if our model adopts a multi-channel approach. Therefore, a model that considers the outlined can be robust against noise, bias field, and TFE and lead to better results on other methods and studies about the brain.

6.2   Hypothesis

We confirmed the stated research hypotheses. The combination of fuzzy methods and adjusted prior knowledge provides better segmentation performance in brain $T1_w$ MRI scans than those models based only on plain clustering.

6.3   Contributions

This research proposed a better approximation to the ill-posed problem of segmentation in brain MRI such that the segmentation parcellates the brain MRI scan into three main tissues, named Cerebrospinal Fluid, Gray Matter, and White Matter.

The main contributions of this research are two; they are in terms of the specific objectives settled. The contributions are:

1. A procedure to fit a generic brain atlas to a specific target. We achieved this goal through the proposed Gardens2 algorithm.

2. An extension of the SOM algorithm, named PLA-SOM, where the topological information of the map and prior knowledge participate in the mapping phase.

3. A robust brain tissue segmentation pipeline.

6.4 Limitations

So far, we have found that the main constraints in this work are:

- The proposed method is limited to the use of $T1_w$, this approach limits its capability to work with MRI sequences with low image resolution and high noise presence.

- The parameter adjustment for Gardens2 and PLA-SOM follows a manual strategy based on experimental results. An automatic optimization method is required so that these parameters adjust to different databases.

- Gardens2 algorithm assumes the availability of a generic brain atlas template.

- Multi-atlas architectures, as presented in Devi et al. (2015); Rajchl et al. (2016), were not explored in our analysis.

- In its current version, Gardens2 and PLA-SOM can be applied as an indirect aid to assess pathologies based on brain volume reduction (progressive brain atrophy). Other types of pathologies, such as lesions provoked by brain tumors El-Dahshan et al. (2014); Subudhi et al. (2018), Multiple Sclerosis Shiee et al. (2010), or Alzheimer's disease Prasad et al. (2015), are ruled out of this proposed method because the analysis of such neurological disorders requires MRI scans different to $T1_w$.

6.5 Future work

As future work, we propose to address the next lines of research:

- To explore a multi-channel approach. At present, a standard MRI study comprises scans in $T1_w$, $T2_w$, and PD_w modalities. Combining this data can provide rich information content to overcome the physics limitations such as low resolution or severe bias field image conditions. As a first step, we will explore the formulation of a regularization parameter based on each voxel's relationship on different MRI sequences.

- To expand the major tissue segmentation into substructure parcellation. The substructure approach can perform the same task that major segmentation approach, but it provides more detailed information about the human brain's changes. For this task, it is necessary to include a topological atlas. As a first step, we will replace the ICBM 2009a Nonlinear atlas with the Individual Adult Brain Atlases[1], which contain a parcellation of the brain into 95 subregions.

- To explore the detection of pathological tissues. One of the main concerns around the study of the human brain is the detection of pathological tissue. This point is the most bold of future work. To adequately address it, the participation of an expert in the clinical area is essential.

6.6   Final remarks

This investigation could impact computational developments and clinical applications. The novel mapping approach presented in PLA-SOM is useful for dealing with the uncertainty in brain tissue segmentation. However, it will be interesting to replace the atlas' information for other types of prior data to apply PLA-SOM in other domains beyond clinical. On the other hand, Gardens2 and PLA-SOM can be extended to work in other clinical applications such as segmentation of spine, lunge, and pathological tissues.

---

[1] http://brain-development.org/brain-atlases/adult-brain-atlases/individual-adult-brain-atlases-new/

References

Abdelsamea, M.M., Gnecco, G., Gaber, M.M., 2015. An efficient Self-Organizing Active Contour model for image segmentation. Neurocomputing 149, 820–835.

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., 2012. Slic superpixels compared to state-of-the-art superpixel methods. IEEE transactions on pattern analysis and machine intelligence 34, 2274–2282.

Agnello, L., Comelli, A., Ardizzone, E., Vitabile, S., 2016. Unsupervised tissue classification of brain MR images for voxel-based morphometry analysis. International Journal of Imaging Systems and Technology 26, 136–150.

Ahmadvand, A., Daliri, M.R., 2015. Improving the runtime of MRF based method for MRI brain segmentation. Applied Mathematics and Computation 256, 808–818.

Al-Dmour, H., Al-Ani, A., 2018. A clustering fusion technique for mr brain tissue segmentation. Neurocomputing 275, 546–559.

Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, J.V., Rueckert, D., 2009. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. Neuroimage 46, 726–738.

Amiri, S., Movahedi, M.M., Kazemi, K., Parsaei, H., 2017. 3d cerebral MR image segmentation using multiple-classifier system. Medical & Biological engineering & Computing 55, 353–364.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26, 839–851.

Aubert-Broche, B., Griffin, M., Pike, G.B., Evans, A.C., Collins, D.L., 2006. Twenty new digital brain phantoms for creation of validation image data bases. IEEE transactions on medical imaging 25, 1410–1416.

Balafar, M.A., Ramli, A.R., Saripan, M.I., Mashohor, S., 2010. Review of brain mri image segmentation methods. Artificial Intelligence Review 33, 261–274.

Bezdek, J.C., Hall, L., Clarke, L., et al., 1993. Review of mr image segmentation techniques using pattern recognition. Medical physics-lancaster 20, 1033– 1033.

Brett, M., Johnsrude, I.S., Owen, A.M., 2002. The problem of functional localization in the human brain. Nature reviews neuroscience 3, 243–249.

Bricq, S., Collet, C., Armspach, J.P., 2008. Unifying framework for multimodal brain MRI segmentation based on hidden Markov chains. Medical image analysis 12, 639–652.

Buxton, R.B., 2009. Introduction to functional magnetic resonance imaging: principles and techniques. Cambridge university press.

Cabezas, M., Oliver, A., Lladó, X., Freixenet, J., Cuadra, M.B., 2011. A review of atlas-based segmentation for magnetic resonance brain images. Computer methods and programs in biomedicine 104, e158–e177.

Caldairou, B., Rousseau, F., Passat, N., Habas, P., Studholme, C., Heinrich, C., 2009. A non-local fuzzy segmentation method: Application to brain MRI, in: International Conference on Computer Analysis of Images and Patterns, Springer. pp. 606–613.

Chen, C., Xie, W., Huang, W., Rong, Y., Ding, X., Huang, Y., Xu, T., Huang, J., 2019. Progressive feature alignment for unsupervised domain adaptation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 627–636.

Chen, S., Zhang, D., 2004. Robust image segmentation using fcm with spatial constraints based on new kernel-induced distance measure. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 34, 1907–1916.

Choi, H.S., Haynor, D.R., Kim, Y., 1991. Partial volume tissue classification of multichannel magnetic resonance images-a mixel model. IEEE Transactions on Medical Imaging 10, 395–407.

Choudhry, M.S., Kapoor, R., 2016. Performance analysis of fuzzy c-means clustering methods for MRI image segmentation. Procedia Computer Science 89, 749–758.

Cicek, S., Soatto, S., 2019. Unsupervised domain adaptation via regularized conditional alignment, in: Proceedings of the IEEE International Conference on Computer Vision, pp. 1416–1425.

Cohen, J., Hillsdale, N., 1988. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates.

Delgado, R.S.F., 2017. Meta-aprendizaje auto-organizado temporal para predicción a largo plazo de series de tiempo caóticas. Ph.D. thesis. Instituto Nacional de Astrofísica, Óptica y Electrónica.

Demirhan, A., Güler, İ., 2011. Combining stationary wavelet transform and self-organizing maps for brain MR image segmentation. Engineering Applications of Artificial Intelligence 24, 358–367.

Deserno, T.M., 2010. Fundamentals of biomedical image processing, in: Biomedical Image Processing. Springer, pp. 1–51.

Despotović, I., Goossens, B., Philips, W., 2015. MRI segmentation of the human brain: challenges, methods, and applications. Computational and mathematical methods in medicine 2015.

Devi, C.N., Chandrasekharan, A., Sundararaman, V., Alex, Z.C., 2015. Neonatal brain MRI segmentation: A review. Computers in biology and medicine 64, 163–178.

Dong, F., Peng, J., 2014. Brain mr image segmentation based on local gaussian mixture model and nonlocal spatial regularization. Journal of Visual Communication and Image Representation 25, 827–839.

El-Dahshan, E.S.A., Mohsen, H.M., Revett, K., Salem, A.B.M., 2014. Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm. Expert systems with Applications 41, 5526–5545.

Engelbrecht, A.P., 2007. Computational intelligence: an introduction. John Wiley & Sons.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., et al., 2012. 3d slicer as an image computing platform for the quantitative imaging network. Magnetic resonance imaging 30, 1323–1341.

Fonov, V., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., Collins, D.L., Group, B.D.C., et al., 2011. Unbiased average age-appropriate atlases for pediatric studies. Neuroimage 54, 313–327.

Friedman, M., 1940. A comparison of alternative tests of significance for the problem of m rankings. The Annals of Mathematical Statistics 11, 86–92.

Girard, G., Whittingstall, K., Deriche, R., Descoteaux, M., 2014. Towards quantitative connectivity analysis: reducing tractography biases. Neuroimage 98, 266–278.

Gonzalez, R.C., Woods, R.E., Eddins, S.L., 2004. Digital image processing using MATLAB. Pearson Education India.

Grande-Barreto, J., Gómez-Gil, P., 2018. Unsupervised brain tissue segmentation in MRI images, in: 2018 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), IEEE. pp. 1–6.

Grande-Barreto, J., Gómez-Gil, P., 2020. Segmentation of mri brain scans using spatial constraints and 3d features. Medical & Biological Engineering & Computing , 1–12.

Grupen, C., Buvat, I., 2011. Handbook of particle detection and imaging. Springer Science & Business Media.

Han, X., Fischl, B., 2007. Atlas renormalization for improved brain mr image segmentation across scanner platforms. IEEE transactions on medical imaging 26, 479–486.

Haralick, R.M., Shanmugam, K., Dinstein, I.H., 1973. Textural features for image classification. IEEE Transactions on systems, man, and cybernetics , 610–621.

Haykin, S.S., Haykin, S.S., Haykin, S.S., Haykin, S.S., 2009. Neural networks and learning machines. volume 3. Pearson Upper Saddle River, NJ, USA:.

IBSR, 2007. Internet brain segmentation repository. Massachusetts General Hospital, Center for Morphometric Analysis. URL: https://www.nitrc.org/projects/ibsr.

Jain, A.K., Duin, R.P.W., Mao, J., 2000. Statistical pattern recognition: A review. IEEE Transactions on pattern analysis and machine intelligence 22, 4–37.

Jain, A.K., Murty, M.N., Flynn, P.J., 1999. Data clustering: a review. ACM computing surveys (CSUR) 31, 264–323.

Ji, Z., Sun, Q., 2017. A fuzzy clustering with bounded spatial probability for image segmentation, in: 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), IEEE. pp. 1–6.

Jiang, Y., Zhao, K., Xia, K., Xue, J., Zhou, L., Ding, Y., Qian, P., 2019. A novel distributed multitask fuzzy clustering algorithm for automatic mr brain image segmentation. Journal of medical systems 43, 118.

Johnson, H., Harris, G., Williams, K., et al., 2007. Brainsfit: mutual information rigid registrations of whole-brain 3D images. Insight J 57, 1–10.

Kendall, M., Gibbons, J., 1990. Rank correlation methods, trans. JD Gibbons (5th edn ed.). Edward Arnold: London .

Kennedy, D.N., Haselgrove, C., Hodge, S.M., Rane, P.S., Makris, N., Frazier, J.A., 2012. Candishare: a resource for pediatric neuroimaging data.

Kohonen, T., 1990. The self-organizing map. Proceedings of the IEEE 78, 1464–1480.

Kong, Y., Chen, X., Wu, J., Zhang, P., Chen, Y., Shu, H., 2018. Automatic brain tissue segmentation based on graph filter. BMC medical imaging 18, 9.

Lee, D.H., 2013. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks, in: Workshop on challenges in representation learning, International Conference on Machine Learning (ICML), p. 2.

Li, S.Z., 2009. Markov random field modeling in image analysis. Springer Science & Business Media.

Lin, G.C., Wang, W.J., Kang, C.C., Wang, C.M., 2012. Multispectral MR images segmentation based on fuzzy knowledge and modified seeded region growing. Magnetic resonance imaging 30, 230–246.

Liu, H., Xu, L., Song, E., Jin, R., Hung, C.C., 2019. Automatic labeling of mr brain images through the hashing retrieval based atlas forest. Journal of medical systems 43, 241.

Liu, J., Zhang, H., 2013. Image segmentation using a local GMM in a variational framework. Journal of mathematical imaging and vision 46, 161–176.

Liu, L., Wang, B., Zhang, L., 2010. An approach based on self-organizing map and fuzzy membership for decomposition of mixed pixels in hyperspectral imagery. Pattern recognition letters 31, 1388–1395.

Lötjönen, J.M., Wolz, R., Koikkalainen, J.R., Thurfjell, L., Waldemar, G., Soininen, H., Rueckert, D., Initiative, A.D.N., et al., 2010. Fast and robust multi-atlas segmentation of brain magnetic resonance images. Neuroimage 49, 2352–2365.

Maaten, L.v.d., Hinton, G., 2008. Visualizing data using t-SNE. Journal of machine learning research 9, 2579–2605.

Magnotta, V.A., Harris, G., Andreasen, N.C., O'Leary, D.S., Yuh, W.T., Heckel, D., 2002. Structural MR image processing using the BRAINS2 toolbox. Computerized Medical Imaging and Graphics 26, 251–264.

Maillard, P., Delcroix, N., Crivello, F., Dufouil, C., Gicquel, S., Joliot, M., Tzourio-Mazoyer, N., Alpérovitch, A., Tzourio, C., Mazoyer, B., 2008. An automated procedure for the assessment of white matter hyperintensities

by multispectral (t1, t2, pd) mri and an evaluation of its between-centre reproducibility based on two large community databases. Neuroradiology 50, 31–42.

Makropoulos, A., Counsell, S.J., Rueckert, D., 2018. A review on automatic fetal and neonatal brain mri segmentation. NeuroImage 170, 231–248.

Mayer, A., Greenspan, H., 2009. An adaptive mean-shift framework for mri brain segmentation. IEEE transactions on medical imaging 28, 1238–1250.

Merényi, E., Taylor, J., 2019. Empowering graph segmentation methods with soms and conn similarity for clustering large and complex data. Neural Computing and Applications , 1–18.

Meyer, F., Beucher, S., 1990. Morphological segmentation. Journal of visual communication and image representation 1, 21–46.

Moon, T.K., 1996. The expectation-maximization algorithm. IEEE Signal processing magazine 13, 47–60.

Negnevitsky, M., 2005. Artificial intelligence: a guide to intelligent systems. Pearson education.

Noe, A., Gee, J.C., 2001. Partial volume segmentation of cerebral MRI scans with mixture model clustering, in: Information Processing in Medical Imaging, Springer. pp. 423–430.

Nordenskjöld, R., 2014. Analysis of Human Brain MRI: Contributions to Regional Volume Studies. Ph.D. thesis. Acta Universitatis Upsaliensis.

Ortiz, A., Górriz, J., Ramírez, J., Salas-Gonzalez, D., Llamas-Elvira, J.M., 2013a. Two fully-unsupervised methods for MR brain image segmentation using SOM-based strategies. Applied Soft Computing 13, 2668–2682.

Ortiz, A., Palacio, A.A., Górriz, J.M., Ramírez, J., Salas-González, D., 2013b. Segmentation of brain MRI using SOM-FCM based method and 3D statistical descriptors. Computational and mathematical methods in medicine 2013.

Pereira, S., Pinto, A., Oliveira, J., Mendrik, A.M., Correia, J.H., Silva, C.A., 2016. Automatic brain tissue segmentation in MR images using random forests and conditional random fields. Journal of neuroscience methods 270, 111–123.

Pham, D.L., 2001a. Robust fuzzy segmentation of magnetic resonance images, in: Computer-Based Medical Systems, 2001. CBMS 2001. Proceedings. 14th IEEE Symposium on, IEEE. pp. 127–131.

Pham, D.L., 2001b. Spatial models for fuzzy clustering. Computer vision and image understanding 84, 285–297.

Pham, D.L., Xu, C., Prince, J.L., 2000. Current methods in medical image segmentation. Annual review of biomedical engineering 2, 315–337.

Porz, N., Bauer, S., Pica, A., Schucht, P., Beck, J., Verma, R.K., Slotboom, J., Reyes, M., Wiest, R., 2014. Multi-modal glioblastoma segmentation: man versus machine. PloS one 9, e96873.

Prakash, R.M., Kumari, R.S.S., 2017. Spatial fuzzy c means and expectation maximization algorithms with bias correction for segmentation of mr brain images. Journal of medical systems 41, 15.

Prasad, G., Joshi, S.H., Nir, T.M., Toga, A.W., Thompson, P.M., (ADNI, A.D.N.I., et al., 2015. Brain connectivity and novel network measures for alzheimer's disease classification. Neurobiology of aging 36, S121–S131.

Rajchl, M., Baxter, J.S., McLeod, A.J., Yuan, J., Qiu, W., Peters, T.M., Khan, A.R., 2016. Hierarchical max-flow segmentation framework for multi-atlas segmentation with kohonen self-organizing map based gaussian mixture modeling. Medical image analysis 27, 45–56.

Rohlfing, T., 2011. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable. IEEE transactions on medical imaging 31, 153–163.

Saikumar, T., Yugander, P., Murthy, P., Smitha, B., 2012. Improved fuzzy c-means clustering algorithm using watershed transform on level set method for image segmentation. International Journal of Machine Learning and Computing 2, 19.

Saito, K., Ushiku, Y., Harada, T., 2017. Asymmetric tri-training for unsupervised domain adaptation, in: Proceedings of the 34th International Conference on Machine Learning-Volume 70, Journal of Machine Learning Research. pp. 2988–2997.

Sandryhaila, A., Moura, J.M., 2013. Discrete signal processing on graphs: Graph filters, in: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE. pp. 6163–6166.

Shattuck, D.W., Sandor-Leahy, S.R., Schaper, K.A., Rottenberg, D.A., Leahy, R.M., 2001. Magnetic resonance image tissue classification using a partial volume model. NeuroImage 13, 856–876.

Shenoy, R., Shih, M.C., Rose, K., 2016. Deformable registration of biomedical images using 2D hidden markov models. IEEE Transactions on Image Processing 25, 4631–4640.

Shiee, N., Bazin, P.L., Ozturk, A., Reich, D.S., Calabresi, P.A., Pham, D.L., 2010. A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. NeuroImage 49, 1524–1535.

Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE transactions on medical imaging 17, 87–97.

Smith, S.M., 2002. Fast robust automated brain extraction. Human brain mapping 17, 143–155.

Soret, M., Bacharach, S.L., Buvat, I., 2007. Partial-volume effect in PET tumor imaging. Journal of Nuclear Medicine 48, 932–945.

Su, M.C., Liu, T.K., Chang, H.T., 1999. An efficient initialization scheme for the self-organizing feature map algorithm, in: IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339), IEEE. pp. 1906–1910.

Subudhi, A., Jena, S., Sabut, S., 2018. Delineation of the ischemic stroke lesion based on watershed and relative fuzzy connectedness in brain mri. Medical & biological engineering & computing 56, 795–807.

Taha, A.A., Hanbury, A., 2015. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. BMC medical imaging 15, 29.

Tesař, L., Shimizu, A., Smutek, D., Kobatake, H., Nawano, S., 2008. Medical image analysis of 3d ct images based on extension of haralick texture features. Computerized Medical Imaging and Graphics 32, 513–520.

Thevenot, J., Chen, J., Finnilä, M., Nieminen, M., Lehenkari, P., Saarakkala, S., Pietikäinen, M., 2014. Local binary patterns to evaluate trabecular bone structure from micro-CT data: application to studies of human osteoarthritis, in: European Conference on Computer Vision, Springer. pp. 63–79.

Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4itk: improved n3 bias correction. IEEE transactions on medical imaging 29, 1310–1320.

Valverde, S., Oliver, A., Cabezas, M., Roura, E., Lladó, X., 2015. Comparison of 10 brain tissue segmentation methods using revisited IBSR annotations. Journal of Magnetic Resonance Imaging 41, 93–101.

Valverde, S., Oliver, A., Roura, E., González-Villà, S., Pareto, D., Vilanova, J.C., Ramió-Torrentà, L., Rovira, À., Lladó, X., 2017. Automated tissue segmentation of MR brain images in the presence of white matter lesions. Medical image analysis 35, 446–457.

Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P., 2003. A unifying framework for partial volume segmentation of brain MR images. IEEE Transactions on medical imaging 22, 105–119.

Verma, H., Agrawal, R.K., Kumar, N., 2014. Improved fuzzy entropy clustering algorithm for MRI brain image segmentation. International Journal of Imaging Systems and Technology 24, 277–283.

Vrooman, H.A., Cocosco, C.A., van der Lijn, F., Stokking, R., Ikram, M.A., Vernooij, M.W., Breteler, M.M., Niessen, W.J., 2007. Multi-spectral brain tissue segmentation using automatically trained k-Nearest-Neighbor classification. Neuroimage 37, 71–81.

Warfield, S.K., Kaus, M., Jolesz, F.A., Kikinis, R., 2000. Adaptive, template moderated, spatially varying statistical classification. Medical image analysis 4, 43–55.

Wilcoxon, F., 1992. Individual comparisons by ranking methods, in: Breakthroughs in statistics. Springer, pp. 196–202.

Wolz, R., Aljabar, P., Rueckert, D., Heckemann, R.A., Hammers, A., 2009. Segmentation of subcortical structures and the hippocampus in brain mri using graph-cuts and subject-specific a-priori information, in: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE. pp. 470–473.

Wu, D., Ma, T., Ceritoglu, C., Li, Y., Chotiyanonta, J., Hou, Z., Hsu, J., Xu, X., Brown, T., Miller, M.I., et al., 2016. Resource atlases for multi-atlas brain segmentations with multiple ontology levels based on t1-weighted mri. Neuroimage 125, 120–130.

Xu, D.H., Kurani, A.S., Furst, J.D., Raicu, D.S., 2004. Run-length encoding for volumetric texture. Heart 27, 452–458.

Zhang, J., Jiang, W., 2014. Segmentation for brain magnetic resonance images using dual-tree complex wavelet transform and spatial constrained self-organizing tree map. International journal of imaging systems and technology 24, 208–214.

Zhang, J., Jiang, W., Wang, R., Wang, L., 2014. Brain MR image segmentation with spatial constrained k-mean algorithm and dual-tree complex wavelet transform. Journal of medical systems 38, 93.

Zhang, X., Wang, G., Su, Q., Guo, Q., Zhang, C., Chen, B., 2017. An improved fuzzy algorithm for image segmentation using peak detection, spatial information and reallocation. Soft Computing 21, 2165–2173.

Zhang, Y., Brady, M., 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.

Additional evaluations

A.1 Segmentation results of Gardens2

In this section we present the segmentation results obtained the proposed algorithm Gardens2. The results come from the execution of the segmentation mask described in Section 4.2.7. The datasets are the same described in Section 5.1.

From Table A.1 and 5.6, we can see that Gardens2 reported a remarkable result in terms of mean and standard deviation (0.79 ± 0.04), but lower than the combination of Gardens2-PLA-SOM (0.82 ± 0.04). Gardens2 reported the second best (0.77 ± 0.03) and the third best (0.56 ± 0.03) results in comparison to those reported in Table 5.9. Box plot depicted on Figure A.1 complement descriptive statistic analysis.

Table A.1: Descriptive statistics of Gardens2 on three datasets.

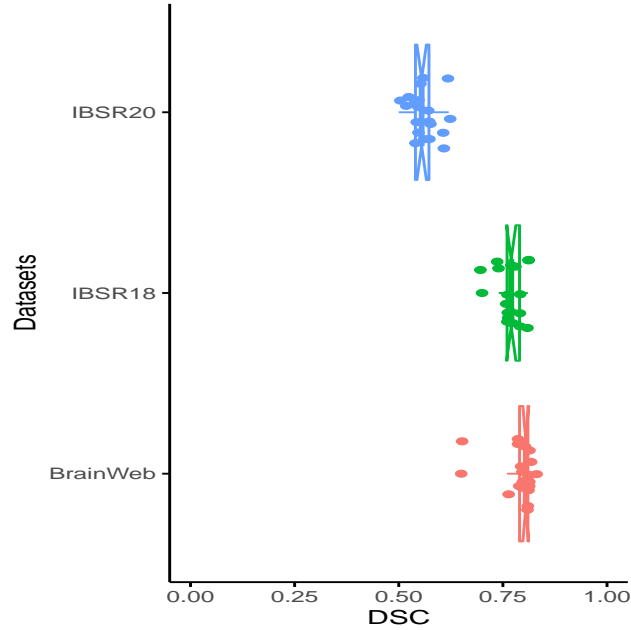| Datasets	Min	Max	Median	Q1	Q3	IQR	Mean
BrainWeb	0.65	0.83	0.81	0.79	0.81	0.02	0.79 ± 0.04
IBSR18	0.70	0.81	0.77	0.76	0.79	0.03	0.77 ± 0.03
IBSR20	0.50	0.62	0.56	0.54	0.57	0.03	0.56 ± 0.03

Figure A.1: Box plot for the DSC from Gardens2 on three datasets.

A.2 The experiment: Performance comparison of PLA-SOM using different pseudo-label sources

This experiment's goal is to assess the performance of PLA-SOM when pseudo labels come from a source different from Gardens2. The experimental hypothesis states that PLA can improve a base method's segmentation performance if the output of that method is constrained fuzzy/probabilistic label vectors.

A.2.1 Results on BrainWeb and IBSR databases

From Table A.2, we observe that PLA-SOM enhanced FANTASM, FAST, FCM_S1, and US performance. The box plot depicted in Figure A.2 shows that PLA-SOM improves the segmentation performance by reducing the dispersion. Table A.3 shows that the best results reported by PLA-SOM on BrainWeb dataset for CSF, GM, and WM correspond to 11% enhancements for the US method, 6% for the FANTASM method, and 4% for the FCMS1 method, respectively. We observe a minimum segmentation improvement, for all the segmentation methods, of 1% for each tissue class on the BrainWeb dataset.
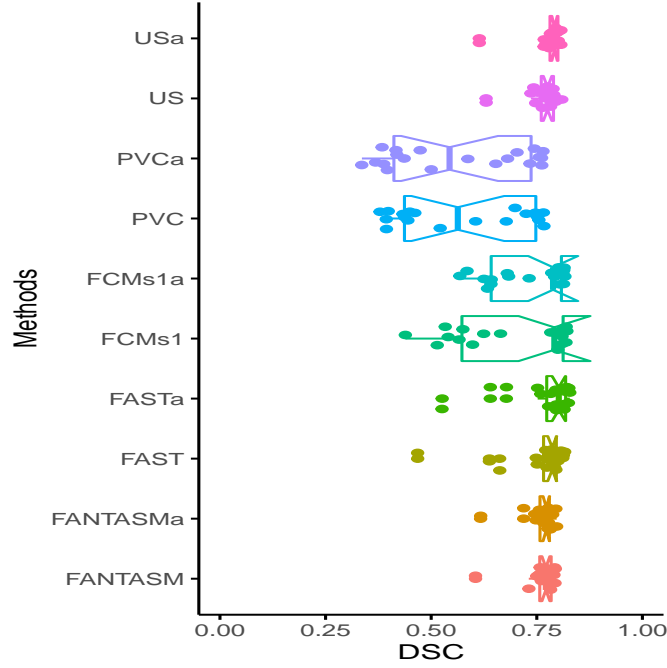
Figure A.2: Box plot for the DSC comparison between single method and the combination with PLA-SOM on the BrainWeb dataset.

Table A.2: Segmentation performance of PLA-SOM on the BrainWeb dataset.

Method	CSF		GM		WM	
	original	PLA-SOM	original	PLA-SOM	original	PLA-SOM
FANTASM	0.75 ± 0.09	0.76 ± 0.09	0.74 ± 0.02	0.79 ± 0.02	0.80 ± 0.02	0.83 ± 0.02
FAST	0.71 ± 0.13	0.76 ± 0.14	0.75 ± 0.07	0.81 ± 0.05	0.86 ± 0.05	0.85 ± 0.04
FCM_S1	0.63 ± 0.20	0.72 ± 0.13	0.71 ± 0.11	0.76 ± 0.08	0.75 ± 0.10	0.79 ± 0.08
PVC	0.48 ± 0.29	0.49 ± 0.30	0.52 ± 0.15	0.54 ± 0.15	0.73 ± 0.05	0.73 ± 0.05
US	0.68 ± 0.09	0.79 ± 0.09	0.78 ± 0.03	0.82 ± 0.03	0.85 ± 0.02	0.86 ± 0.02

Table A.3: Descriptive statistics of PLA-SOM on the BrainWeb dataset.

Method	Min		Max		Median		IQR		Mean	
	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM
FANTASM	0.60	0.64	0.79	0.83	0.77	0.80	0.03	0.03	0.76 ± 0.04	0.79 ± 0.04
FAST	0.47	0.53	0.82	0.85	0.79	0.84	0.03	0.02	0.76 ± 0.08	0.81 ± 0.08
FCM_S1	0.44	0.59	0.82	0.85	0.79	0.82	0.24	0.17	0.70 ± 0.13	0.76 ± 0.09
PVC	0.38	0.34	0.77	0.78	0.56	0.57	0.31	0.32	0.58 ± 0.16	0.58 ± 0.16
US	0.63	0.64	0.81	0.85	0.78	0.83	0.03	0.01	0.77 ± 0.04	0.82 ± 0.04

Table A.4: Results obtained on IBSR18 and IBSR20 datasets.

IBSR18						
Method	CSF		GM		WM	
	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM
FANTASM	0.45 ± 0.11	0.51 ± 0.11	0.72 ± 0.07	0.77 ± 0.06	0.88 ± 0.03	0.87 ± 0.02
FAST	0.57 ± 0.06	0.65 ± 0.06	0.80 ± 0.05	0.84 ± 0.03	0.89 ± 0.02	0.87 ± 0.01
FCM_S1	0.44 ± 0.12	0.49 ± 0.12	0.70 ± 0.08	0.75 ± 0.07	0.73 ± 0.06	0.85 ± 0.02
PVC	0.58 ± 0.07	0.67 ± 0.09	0.75 ± 0.08	0.79 ± 0.06	0.85 ± 0.07	0.83 ± 0.05
US	0.48 ± 0.09	0.63 ± 0.06	0.82 ± 0.03	0.84 ± 0.03	0.86 ± 0.01	0.85 ± 0.02

IBSR20						
Method	CSF		GM		WM	
	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM
FANTASM	0.20 ± 0.07	0.21 ± 0.07	0.71 ± 0.06	0.73 ± 0.06	0.79 ± 0.05	0.80 ± 0.05
FAST	0.18 ± 0.05	0.21 ± 0.06	0.70 ± 0.04	0.75 ± 0.04	0.80 ± 0.06	0.80 ± 0.05
FCM_S1	0.20 ± 0.07	0.20 ± 0.07	0.70 ± 0.04	0.72 ± 0.03	0.77 ± 0.03	0.78 ± 0.07
PVC	0.17 ± 0.06	0.21 ± 0.07	0.68 ± 0.07	0.74 ± 0.06	0.74 ± 0.19	0.73 ± 0.21
US	0.24 ± 0.06	0.25 ± 0.06	0.81 ± 0.03	0.82 ± 0.02	0.83 ± 0.03	0.82 ± 0.03

Table A.5: Descriptive statistics of PLA-SOM on IBSR18 and IBSR20 datasets.

IBSR18										
Method	Min		Max		Median		IQR		Mean	
	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM
FANTASM	0.58	0.63	0.78	0.82	0.69	0.72	0.06	0.06	0.68 ± 0.06	0.72 ± 0.05
FAST	0.70	0.73	0.80	0.83	0.76	0.79	0.03	0.03	0.75 ± 0.03	0.79 ± 0.03
FCM_S1	0.54	0.56	0.83	0.81	0.66	0.70	0.06	0.07	0.67 ± 0.06	0.70 ± 0.06
PVC	0.63	0.64	0.83	0.85	0.74	0.77	0.10	0.08	0.73 ± 0.06	0.76 ± 0.06
US	0.67	0.71	0.77	0.81	0.72	0.78	0.04	0.03	0.72 ± 0.03	0.77 ± 0.02

IBSR20										
Method	Min		Max		Median		IQR		Mean	
	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM	Original	PLA-SOM
FANTASM	0.45	0.47	0.65	0.66	0.57	0.58	0.04	0.04	0.56 ± 0.05	0.58 ± 0.05
FAST	0.42	0.46	0.61	0.65	0.56	0.59	0.03	0.04	0.56 ± 0.04	0.58 ± 0.04
FCM_S1	0.42	0.44	0.62	0.63	0.55	0.57	0.03	0.02	0.56 ± 0.04	0.58 ± 0.04
Gardens2	0.50	0.47	0.62	0.65	0.55	0.58	0.03	0.05	0.56 ± 0.03	0.57 ± 0.04
PVC	0.29	0.30	0.63	0.66	0.56	0.57	0.06	0.06	0.53 ± 0.09	0.56 ± 0.09
US	0.60	0.60	0.66	0.66	0.63	0.63	0.03	0.03	0.63 ± 0.02	0.63 ± 0.02

Table A.4, the best results of PLA-SOM on the IBSR18 dataset for CSF, GM, and WM are 15% for the US method, 5% for the FANTASM method, and 12% for the FCM_S1 method, respectively. There is an improvement of 2% in the
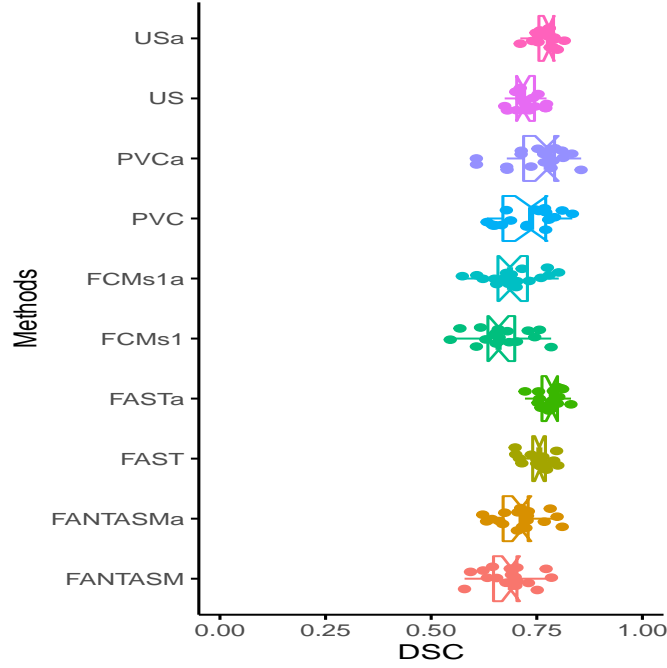
Figure A.3: Box plot for the DSC comparison between single method and the combination with PLA-SOM on the IBSR18 dataset.

segmentation performance on the CSF and GM classes. Figure A.3 shows that PLA-SOM reduced the segmentation performance of different methods for the WM class. These results are associated with the low correspondence between the probabilistic atlas and the WM target subject. From Table A.5, we can see that PLA-SOM improved the performance of all the methods. On the other hand, from Table A.4 and A.5, we observe a minimum segmentation improvement for all the segmentation methods of 1% for each tissue class and 2% on mean performance the IBSR20 dataset.

The results presented in this section show evidence that Gardes2 can also achieve results comparable to those reported by the state-of-the-art. On the other hand, PLA-SOM, used as a post-procedure, can improve a base method's segmentation performance. An important characteristic of both algorithms is their modularity and their ability to be combined with other methods that address the same problem of brain tissue segmentation.
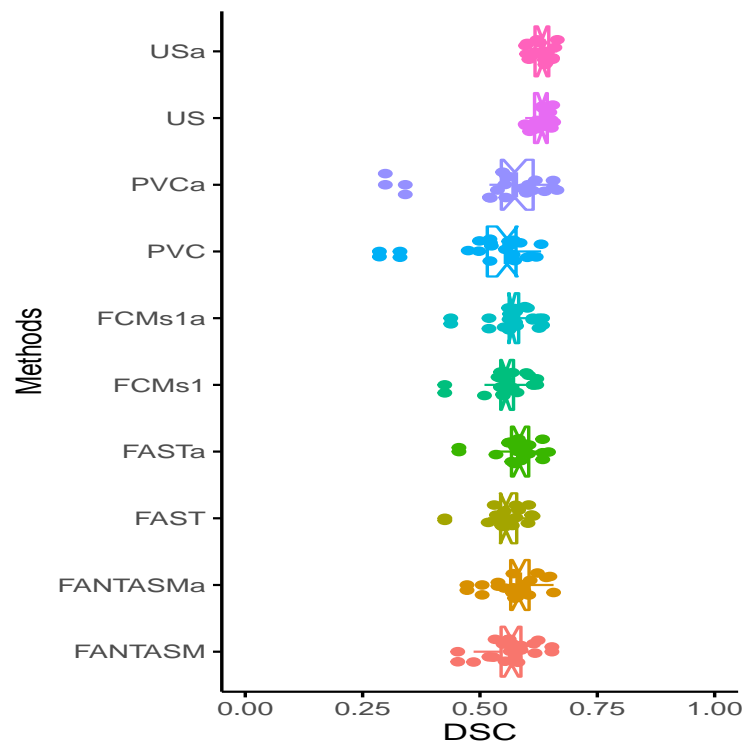
Figure A.4: Box plot for the DSC comparison between single method and the combination with PLA-SOM on the IBSR20 dataset.