UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Departamento de Sistemas Informáticos y Computación
Máster en Inteligencia Artificial, Reconocimiento de
Formas e Imagen Digital

Unsupervised glioblastoma segmentation based on multiparametric Magnetic Resonance Imaging (MRI)

M.Sc. Thesis

Presented by **Javier Juan Albarracín**

Supervised by Dr. Alfons Juan i Císcar
and Dr. Juan Miguel García Gómez
and Dr. Elies Fuster i Garcia

Valencia - Spain
September 11, 2014

Agradecimientos

La finalización de esta tesina de máster no habría sido posible sin el apoyo de un conjunto de personas que me han acompañado durante todo este trayecto.

Quiero comenzar por mi más sincero agradecimiento a mis directores de tesis, en especial al Dr. Juan Miguel García y al Dr. Elies Fuster. Gracias por apostar por mí y por ofrecerme un hueco en el grupo de investigación IBIME. Durante este tiempo, he tenido la oportunidad de aprender y desarrollarme como persona y como profesional al lado de un equipo humano increible, al que considero referentes tanto a nivel profesional como personal. Gracias también a Alfons Juan por su amabilidad y por acceder en el último momento a ser director de esta tesis, que de otra manera no habría podido culminarse.

Agradezco tambien a todos los compañeros del grupo IBIME, especialmente a los de la linea de minería de datos y pattern recognition. Ademas de ser excelentes profesionales son tambien excelentes personas de las cuales he aprendido mucho. Gracias al Dr. Salvador Tortajada, Carlos Sáez, Miguel Esparza y Adrián Bresó. Quiero tambien recordar a Alfonso Pérez y al Dr. Javier Vicente, ex-compañeros del grupo, con los cuales he compartido muy buenos momentos. Les deseo el mejor futuro en sus nuevos proyectos.

Esta tesina ha servido como piloto y punto de partida para varios proyectos tanto nacionales como privados, principalmente en consorcio con el Hospital Universitario y Politécnico La Fe y el Grupo Hospitalario Quirón. Quiero agradecer al personal de estas dos instituciones, en concreto a Roberto Sanz Requena, Fernando Aparici y Luis Marti-Bonmati, por la orientación y los acertados consejos y comentarios que me han brindado durante el desarrollo de este trabajo.

Quiero dedicar esta tesina a mi padre. Siempre me has inculcado los valores de lucha y esfuerzo para conseguir los objetivos que uno se plantea. En los momentos en los que la vida te da la espalda es cuando más necesario se hace luchar para superar las adversidades. Ahora es uno de esos momentos, y no estas solo. Estoy seguro de que entre todos lo conseguiremos.

Por último, gracias Itziar por compartir la vida conmigo y hacerla cada dia más fácil. Sin el apoyo y el cariño que me regalas siempre, esto no habría sido posible.

Abstract

Over the years, the medical field has been experimented an evolution in their clinical practices towards the inclusion of new technologies that can assist in the diagnosis and prognosis of complex diseases. Such evolution has derived in the development of Clinical Decision Support Systems (CDSSs) that provide physicians advanced tools to improve their medical aid, decision making and monitoring process of the patients. Rather than replacing the clinician figure, CDSSs are aimed to assist the human to overcome their natural limitations in the analysis of complex and large volumes of information, such as the patient clinical records. CDSSs are defined as computational systems that provide precise and specific knowledge for the medical decisions to be adopted for diagnosis, prognosis, treatment and management of patients. Such definition links the nature of CDSSs to a specific concept: Artificial Intelligence (AI) in medicine.

Brain tumour diagnosis concerns a concrete pathology that has received a lot of attention from the Pattern Recognition (PR) and Machine Learning (ML) community. Due to the heterogeneity and complexity of the different tumours and the huge amount of information handled by the multidisciplinary clinicians groups, CDSSs have become a key component for the future of brain tumour treatment. As a major step in the treatment of brain tumours, the early identification and delineation of the different tissues related to the lesion becomes crucial to make decisions that can improve the patient survivability. In this sense, automatic brain tumour segmentation plays a key role in the development of CDSSs.

Currently, most of the automatic brain tumour segmentation approaches arise from the supervised learning standpoint. The supervised learning paradigm requires a labelled training dataset from which to infer the models of the classes that represent the different tissues in the brain. Such training datasets are usually obtained through expert manual annotations, which becomes a tedious, time-consuming and biased process, among other limitations. On the other hand, unsupervised approaches address these limitations, but usually achieve worse results in comparison to supervised approaches and often require several manual stages to improve the interpretability of their results.

In order to overcome these limitations, this M.Sc. thesis introduces a fully automated unsupervised method for brain tumour segmentation using anatomical Magnetic Resonance (MR) images, able to achieve accurate results comparable with supervised approaches. A PR scheme is adopted to design the fully automated unsupervised segmentation method. First, a preprocessing stage based on the state of the art techniques in MRI is proposed to enhance and correct the information contained in the images. Next, a feature extraction and dimensionality reduction is carried out to extract discriminative features from the images and to simplify the inference process of the unsupervised algorithms. Four unsupervised clustering techniques, divided by their structured or non-

structured condition, are evaluated to assess their pros and cons. Considering the non-structured algorithms, K-means, Fuzzy K-means and Gaussian Mixture Model (GMM) clustering are analysed, whereas as structured classification algorithms, theGaussian Hidden Markov Random Field (GHMRF) is evaluated. Finally, an automated tumour classes isolation based on a statistical approach supported by tissue probability maps is proposed to overcome the lack of biological interpretability of the unsupervised results.

The proposed segmentation method is evaluated through the public International BRAin Tumour Segmentation (BRATS) dataset to compare its performance against the state-of-the-art supervised approaches that participate in the challenge. Our results placed the method in the 7th position of the challenge, with a Dice score of 0.72 for the complete tumour subcompartment, which confirms our approach as a viable alternative for Glioblastoma Multiforme (GBM) segmentation.

Contents

Chapter 1

Introduction

1.1 Motivation

Over the years, the medical field has been experimented an evolution in their clinical practices towards the inclusion of new technologies that can assist in the diagnosis and prognosis of complex diseases. However, it was not until the 20th century since the medicine was completely revolutionized with the explosion of the use of medical technology. Several advances developed in these years comprise the electrocardiography (Willem Einthoven, 1903), the electroencephalography (Hans Berger, 1929), the Heart-Lung machine (Dr John Heysham Gibbon, 1953) and MRI (Raymond Vahan Damadian, 1971). These improvements led to a huge increase of the data used to diagnose and treat patients, thereby converting the clinical records on an important document to store the patient's information.

Nowadays, this large amount of information has even increased due to its multidisciplinary origin, introducing new requirements to the management of the patient information during his disease. Such requirements involve the analysis of complex multi-source and often multi-center clinical data and the integration of medical knowledge from different health areas, in order to improve the quality of the treatment. Advanced systems able to assist the human to overcome their natural limitations in the analysis of complex volumes of information are then required. In this sense, Clinical Decision Support Systems (CDSSs) emerged to provide physicians powerful tools to improve their medical aid, decision making and monitoring process of the patients. CDSSs are closely related to Artificial Intelligence (AI) and Machine Learning (ML) disciplines, as they are aimed to provide precise and specific knowledge for the medical decisions to be adopted for diagnosis, prognosis and treatment of patients. In this sense, the Pattern Recognition (PR) and Machine Learning (ML) community has shown a significant interest in the development of CDSSs due to the complex and crucial task that health represents in the everyday human life.

Specifically, brain tumour has received a lot of this attention due to its incidence in the population and its consequences in their life expectancy. Glioblastoma Multiforme (GBM) tumour is the most common and most aggressive malignant tumour [1, 2], which presents heterogeneous lesions consisting of different areas of active tumour, necrosis and edema, all of them exhibiting a high variability related to the aggressiveness of the tumour. Its intra-cranial location and the unspecificity of clinical symptoms [3] makes medical imaging

techniques to play a key role in the GBM treatment. The standard technique for GBM diagnosis concerns the acquisition of several Magnetic Resonance (MR) images to locate and identify the different tissues related to the lesion. However, the manual analysis and delineation of these relevant tissues involves a complex, tedious, time-consuming and biased task. The increasingly consolidated techniques based on PR approaches have been shown to provide automated efficient solutions for routine clinical application [4].

Nonetheless, most of these techniques arise from the supervised learning standpoint. Supervised learning is a ML approach that depends on a manually annotated training datasets to learn discriminative functions used for mapping new examples. The training dataset consists of a set of observed-input and desired-output pairs, used to infer the models that represents these relationships. The most important drawback of the supervised learning concerns the acquisition of the training dataset [5]. The training dataset must be manually created by experts and should contain a sufficiently large set of labelled examples than enables to learn *generalized* models that apply successfully to unseen data. Hence, supervised learning is limited to the quality and size of the training dataset, which requires an expensive, tedious, time-consuming and biased task to compile it. Furthermore, common problems such as the over-fitting of the models or the inability to provide labels for all situations in some problems should be considered [5]. Finally, a well-known problem of brain tumour MR images is the normalization between acquisitions of different patients [4]. Supervised approaches require a recalibration of the models when the data sources are not normalized or experiment changes, for example different hospitals or new MR protocols. Such limitation in combination with the difficult and expensive task of compile manual labelled brain tumour datasets severely affects the viability of supervised approaches for real clinical routine.

In these sense, this M.Sc. Thesis is intended to provide an unsupervised learning methodology for GBM segmentation, able to overcome the limitations concerning the supervised approaches. The proposed methodology has to reach comparable results to supervised segmentations, while addressing the inherent limitations of the unsupervised learning such as the lack of interpretability of the unsupervised partitioning. Furthermore, the method has to provide mechanisms to minimize the initialization problem of clustering algorithms, which may lead to poor local minima solutions. Finally, it is interesting that the unsupervised method takes advantage of the self similarity presented in the images, by employing models that considers dependencies between the data to be segmented.

1.2 Hypothesis

The present M.Sc. Thesis is based on the following hypothesis:

I The unsupervised approach to brain tumour segmentation is a viable approach as it can obtain coherent and accurate results similar than the ones retrieved by manual expert labelling, and also comparable to supervised segmentations, but avoiding the tedious, time-consuming and biased task of manual expert labelling.

1.3 Goals

The general goal of this M.Sc. Thesis is to contribute in the design, development and validation of brain tumour segmentation methods, specifically in the unsupervised learning field, by providing a complete and robust methodology for GBM segmentation.

This main goal is achieved by fulfilling the following specific goals:

- To design a fully automated unsupervised brain tumour segmentation method. Identify the requirements and stages of the methodology and provide a competitive segmentation system comparable to supervised approaches.
- To study and implement the required techniques and algorithms to develop the unsupervised methodology.
- To design and implement mechanisms to improve the biological interpretability of the unsupervised results for GBM segmentation, as they are devoid of semantic meaning.
- To evaluate the proposed unsupervised segmentation method with a real public and reference brain tumour dataset.
- To compare the performance of the unsupervised segmentation method with state-of-the-art supervised segmentation algorithms.

1.4 Contributions

The scientific contribution of this M.Sc. thesis concerns the application of ML techniques, specifically unsupervised learning algorithms, to design and develop a complete automated method for GBM segmentation. This method is not limited to GBM segmentation but can also be applied to other brain tumour cases and to other pathologies such as Multiple Sclerosis. Nevertheless, this M.Sc. Thesis only evaluates the method for GBM segmentation to focus the study. The technological results of this M.Sc. Thesis are compiled and registered by the Universitat Politècnica de València (UPV) and currently the method is under the registration process as a original patent with reference ID 769-PAT/MGM.

This M.Sc. Thesis has served as a prototype and study case to write the National Research Project *Caracterización de firmas biológicas de Glioblastomas mediante modelos no supervisados de predicción estructurada basados en biomarcadores de imagen (TIN2013-43457-R)*, which has been accepted and will be funded by the Ministerio de Economía y Competitividad of Spain. This Thesis has also served to achieve a Private Research Project called *Segmentación no supervisada de Glioblastomas basada en imagen de resonancia magnética multiparamétrica y restricciones espacio/temporales*, which is funded and supported by Hospital Universitario Politécnico La Fe and BRACCO Company[a].

The main contributions of this M.Sc. thesis can be summarized as follows:

Contribution 1: The design and implementation of a robust methodology for unsupervised brain tumour segmentation by means of a definition of a complete automated

[a] `http://imaging.bracco.com/us-en`

method able to achieve accurate results comparable to supervised approaches, avoiding the tedious, time-consuming and biased task of manual expert labelling.

Contribution 2: The design of a postprocessing stage able to robustly differentiate between pathological and non-pathological classes in a brain tumour segmentation. The statistical approach that underlines the method provides a flexible and powerful framework to accurate identify tissues that not correspond to healthy brains. This approach can be also extrapolated to other pathologies besides GBM, as it can detect anomalies not referred to healthy tissues.

Contribution 3: Evaluation of a complete unsupervised segmentation approach with a public real brain tumour dataset, in order to assess the performance of the unsupervised approaches with respect to the state-of-the-art techniques in supervised segmentation. Several unsupervised algorithms are evaluated within the methodology to asses the benefits and limitations of each one, differentiating between structured and non-structured segmentation algorithms.

1.5 Summary of the remaining chapters

Chapter 2 briefly introduces the principles of MRI and the different MRI modalities used in the study. Furthermore, the public dataset used to evaluate and compare the performance of the method against supervised approaches is presented.

Chapter 3 presents the preprocessing stage for the proposed unsupervised segmentation methodology. Several state of the art techniques proposed to correct the most common artefacts of MRI acquisitions are exposed. The preprocessing covers the following artefacts and operations: Denoising, magnetic bias field correction, skull stripping and superresolution.

Chapter 4 introduces the techniques used to extract discriminative features from the MR images to differentiate between the tumoral tissues. In addition to the intensity levels provided by each MR image, texture features are computed for all sequences to distinguish between tumoral tissues. Dimensionality reduction based on Principal Component Analysis (PCA) is proposed to reduce the high-dimensional data space representation, thereby decreasing the complexity in the inference of the algorithms.

Chapter 5 covers the unsupervised clustering algorithms evaluated in the brain tumour segmentation method. Both structured and non-structured classification algorithms are considered. Under the non-structured paradigm; K-means, Fuzzy K-means and GMM unsupervised clustering algorithms are evaluated. Regarding the structured prediction pattern; GHMRF algorithm is proposed. All clustering techniques are postulated in terms of variants of generative mixture models and EM algorithm.

Chapter 6 introduces the proposed automatic tumoral classes isolation designed to improve the lack of interpretability of the unsupervised results. An automatic identification of pathological classes is presented based on a statistical approach supported by tissue probability maps obtained for normal tissues.

Chapter 7 discloses the evaluation of the unsupervised segmentation system. Figures of merit used for the evaluation of the method are presented. Results obtained for the different clustering techniques in combination with the proposed preprocessing and postprocessing stages are described. A ranking comparing the results achieved by the supervised approaches evaluated in the International Image Segmentation Challenge of Medical Image Computing and Computer-Assisted Intervention (MICCAI) Conference and the proposed method is presented.

Chapter 8 discusses the pros and cons of the different algorithms of the preprocessing stage, as well as the results obtained by the different unsupervised clustering algorithms and the impact of the proposed automatic tumoral classes isolation method.

Chapter 7 summarizes the conclusions and explains the future lines of research and development.

Chapter 2

Review of the literature

Glioblastoma Multiforme (GBM) is the most frequent (>50%) and most aggressive malignant tumour of the Central Nervous System (CNS) [1]. GBMs are heterogeneous malignant masses, characterized by hypercellularity, pleomorphism, microvascular proliferation and high necrosis mitotic activity, in which different areas of malignancy grade can coexist [2]. The current standard treatment for GBM concerns surgery, radiotherapy and chemotherapy, with an average global survivability of 15 months and a progression-free survivability of 7 months [6]. Hence, the early identification of the different malignant tissues related to the tumour becomes crucial to make decisions that can improve the patient survivability. The segmentation of the tumoral and peritumoral areas in combination with abnormal tissue classification such as edema or necrosis is determinant to monitor the evolution of the tumour recurrence or shrinkage during therapy.

The standard technique for GBM diagnosis is MRI [3]. MRI is a medical imaging technique used to provide detailed images of the different types of tissues in the brain through a non-invasive process. MRI is able to produce images with different types of tissue contrast, which enables the segmentation and differentiation of the tissues. Due to the heterogeneous nature of brain tumours, specifically in case of the GBM, several MRI sequences are required to diagnose and segment the tumour including all its subregions [7]. The standard anatomical sequences used for the diagnosis are T_1-weighted, T_1-weighted with contrast enhancement (T_{1c}), T_2-weighted and FLuid Attenuated Inversion Recovery (FLAIR).

In clinical practice, the segmentation is performed manually over these sequences by expert radiologists, which becomes a complicated, tedious and time-consuming task, frequently unaffordable to the humans in a reasonable time. Hence, in recent years, the interest of automated segmentation techniques to recognize both the pathological and healthy tissues of the brain has arisen. A previous review of automated brain tumour segmentation methods was done by Angelini *et al.* [8]. However, the rapid advancement in ML techniques applied to radiology [9] has derived in the most important brain tumour segmentation methods after the 2007. Recent extensive reviews that compile most of these techniques have been presented in [10, 4].

Brain tumour segmentation is usually addressed from the supervised learning standpoint. Cai *et al.* [11] and Verma *et al.* [12] created voxel-wise feature vectors from a large number of MRI sequences, including Diffusion Tensor Imaging (DTI), and applied Support Vector Machines (SVM) to segment the tumour and additional subcompartments inside

the lesion area. Ruan *et al.* [13, 14] used kernel class separability for feature selection in a multiparametric MRI set and also used SVM to segment the tumour region. Tayel *et al.* [15] proposed a combination of neural networks and fuzzy logic rules to segment Region Of Interests (ROIs) for brain tumour diagnosis. Jensen *et al.* [16] proposed an initial ROI segmentation based on morphological and thresholding operations and a posterior training stage where four classifiers, comprising a logistic regression, a multi layer perceptron and both fuzzy inference systems, were trained to segment different tissues of the tumour. Lee *et al.* [17] used a combination of SVM and pseudo Conditional Random Fields (CRF) to brain tumour segmentation. They performed the classic training step of the SVM considering the data independent and identically distributed (iid), but introduces spatial constraints similar as the CRFs to the inference process of the SVM to improve the performance of the segmentation. Bauer *et al.* [18] also used SVM in combination with hierarchical CRF to segment both healthy and tumour tissues, in addition to several sub-compartments inside the lesion.

However, supervised learning requires an expensive, time-consuming and biased task to retrieve manual labelled datasets from which to learn the segmentation models, hence limiting the performance of the supervised approach to the quality of the training dataset. Furthermore, common problems such as the over-fitting or the inability to provide labels at all for some classification problems directly affects the supervised paradigm. Moreover, MRI acquisition protocols are not standardized [4], which in combination with the critical and complex problem of brain tumour MRI normalization, limits the supervised models to new cases acquired under the same protocol and conditions than the one used for train the models. This limitation becomes more important as new clinical data tends to come from different sources and is often shared between hospitals.

Unsupervised learning address this problems in a more straightforward way. Unsupervised learning does not require a training dataset from which to learn the pair relations between observation and labels, but directly uses the data to find natural groupings of observations that represent clusters of information. Hence, the unsupervised learning is much less affected by the heterogeneities between patients presented in the MRI datasets, as it can segment each patient independently with its own data.

Although unsupervised learning is able to address these limitations, few research effort have been done in the brain tumour segmentation field. The heterogeneity of the tumours, specifically in the GBMs tumour, makes the segmentation more challenging if no prior knowledge is considered [19]. Additionally, the biological interpretation of the unsupervised results is not as forward as in the supervised approach, thereby requiring additional techniques to improve its comprehensibility.

Anyway, several attempts for brain tissue segmentation have been made under the unsupervised paradigm. The first unsupervised model for tumour segmentation was proposed by Schad *et al.* [20] in 1993. In this paper, texture patterns computed from the gray level intensities of the MR images were used to classify different ROIs into healthy and pathological tissues. Later, Fletcher *et al.* [21] proposed an approach based on fuzzy clustering and domain knowledge for multi-parametric non-enhancing tumour segmentation. Domain knowledge and parenchymal tissue detection was based on heuristics related to geometric shapes and locations, which may not be robust when the high deformation produced by the tumours is presented. Moreover, several assumptions such as prior knowledge about the number of existing tumours or the slice thickness required for the MRI acqui-

sitions introduced several limitations to the method. Nie *et al.* [22] proposed a Gaussian clustering with a spatial accuracy-weighted Hidden Markov Random Fields (HMRF) that allowed them to deal with images at different resolutions without interpolation. Nowadays, advanced reconstruction techniques such as super-resolution enables to work in a high resolution voxel space, minimizing typical problems of interpolation such as partial volume effects. Moreover, no method was provided to differentiate between tumoral classes and normal tissue classes of the brain, so manual identification might be needed. Zhu *et al.* [23] developed a software based on the segmentation method proposed by Zhang *et al.* [24], which performs an EM Gaussian clustering combined whit HMRF's. Zhu *et al.* extended Zhang's approach through a sequence of additionally morphological and thresholding operations to refine the segmentation, however such operations are not fully specified and only overall commented, so the reproducibility of their results is not possible. Vijayakumar *et al.* [25] proposed a method based on Self Organizing Maps (SOMs) to segment tumour, necrosis, cysts, edema and normal tissues using a multi-parametric MRI set. Although the learning process of SOMs is performed in an unsupervised manner, the dataset from which to infer the structure adopted by the Artificial Neural Networks (ANN) should be selected and determined manually, such as a supervised approach. In the Vijayakumar work, 700 pattern observations were chosen, corresponding to 7 different tissues that they assumed to exist in the brain. The correct selection of these prototypes determines the quality of the network, hence converting the process in a supervised *labelling* task.

Furthermore, all the unsupervised approaches proposed above apply their algorithms on its own datasets, making difficult a general comparison of the methods. In the last years, important efforts has been made to provide public brain tumour segmentation datasets to evaluate the performance of the proposed methods and the current state-of-the-art in automated brain tumour segmentation. MICCAI Society promotes this idea and started at 2012 to organize the BRATS Challenge, by providing a public annotated dataset of high-grade and low-grade gliomas.

Chapter 3

Materials

3.1 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a medical imaging technique used to provide internal representations of the human body, crucial for the diagnosis, follow-up prognosis and treatment of complex diseases. This technique was first discovered by the Armenian-American physician, scientist and professor Raymond Damadian, who published in 1971 a paper in the journal *Science* [26] reporting that tumours could be distinguished in vivo from normal tissues by Nuclear Magnetic Resonance (NMR). Although Damadian's initial method was not viable for practical use, he developed in 1972 the world's first magnetic resonance imaging machine. Simultaneously, Paul Lauterbur, extending the work proposed by Herman Carr, published in 1973 the first MR image [27] and one year after the first cross-sectional image of a living mouse [28]. In the late 1970s, the professor Peter Mansfield developed a mathematical technique able to provide MRI scans that took seconds rather than hours to produce clearer images than the Lauterburg and Damadian's images. In August of 1980 the team led by John Mallard obtained the first clinically useful MRI, which identified a primary tumour in the patient's chest, an abnormal liver and secondary cancer in the patient bones. Paul Lauterbur and Peter Mansfield were later awarded in 2003 with the Nobel Prize for their discoveries and advances in MRI.

MRI is based on the magnetic properties of the atomic nuclei, specifically on the spin angular momentum of the hydrogen nucleus (H^+). At a resting natural state, all the hydrogen H^+ nucleus in the human body spin randomly, thus cancelling the angular momentums each other and producing an overall zero spin magnetic momentum value. Under the influence of an external uniform magnetic field B_0, the H^+ nucleus align its spin with the B_0 in a parallel (low energy) or anti-parallel (high energy) state (S_{B_0}). Then, a sequence of Radio Frequency (RF) pulses at the Larmor frequency of the H^+ nucleus is applied to the B_0, exciting the protons and enforcing them to change its spin orientation 90^o with respect to the B_0 direction ($S_{B_{90}}$). After the RF pulses end, the H^+ nucleus begins to return to the S_{B_0} state by recovering the B_0 direction and the excessive spin begins to dephase at a different frequencies regarding to chemical context of the H^+ nucleus. The magnetic coils capture the electric signals produced by the nucleus during their transition, and store them in the so called K-space, in order to later reconstruct the MR image through an inverse Fourier process.

The rate at which the H^+ nuclei realigns with the B_0 field and its dephase rate de-

termines the different contrasts of the images. There are three types of contrast in MR images: T_1-weighted, T_2-weighted and Proton Density (PD) images. The different contrasts are related to the Repetition Time (TR) and the Echo Time (TE) times. TR is the time between successive RF pulses and affects the speed in which H+ protons realigns to the B_0 filed after the RF ends. The TE refers to the time at which the electrical signal induced by the H+ protons is measured in the magnetic coils and concerns the degree of dephasing of the spins of the protons. Figure 3.1 show the relation between the TR and TE and the contrast of the MR images.

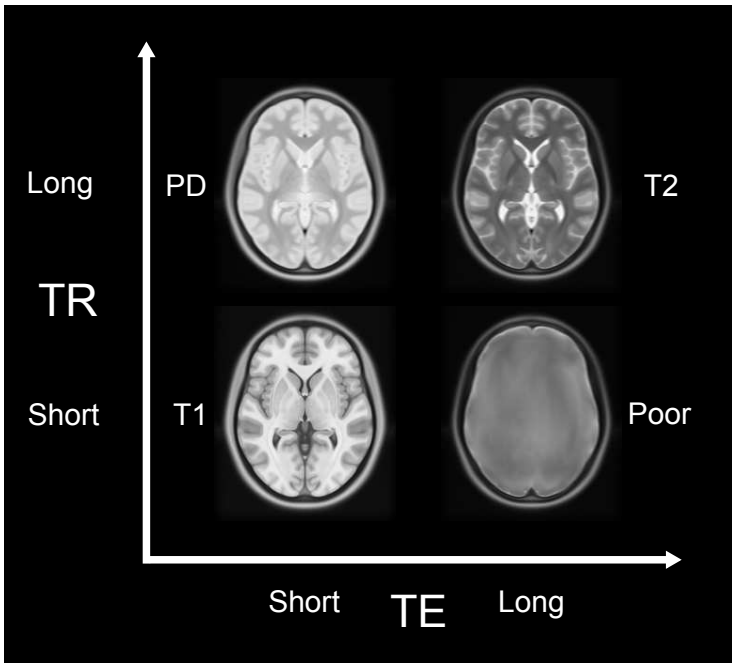

Figure 3.1: Relation between short and long TR and TE and the contrast obtained in MR images.

The clinical gold standard for brain tumour diagnosis relies on the use of T_1 and T_2-weighted sequences, FLuid Attenuated Inversion Recovery (FLAIR) sequence and contrast-enhanced T_1-weighted sequence (T_{1c}) [29, 30]. T_1-weighted sequences are commonly used to differentiate between healthy tissues as they primarily show structural information of the brain. GBM tumours typically present a hypointense to isointense pattern on the lesion area in the T_1-weighted sequence. T_{1c} enhances the tumour borders given that the contrast agent (Gadolinium (Gd)) accumulates in this area due to the disruption of the blood-brain barrier. An enhanced Gadolinium ring around the tumour is typically presented in the T_{1c} sequences. The T_2-weighted images highlights the lesion area including the edema region, presenting a hyperintense pattern in such area. Also the Cerebro-Spinal Fluid (CSF) appears hyperintense, hence avoiding its separation in the T_2 sequence. The FLAIR sequence is a special T_2-weighted image with free water signal suppression, which enables the differentiation between the CSF and the edema. Figure 3.2 shows an example of the visualization of GBM tumour in the different MR sequences.

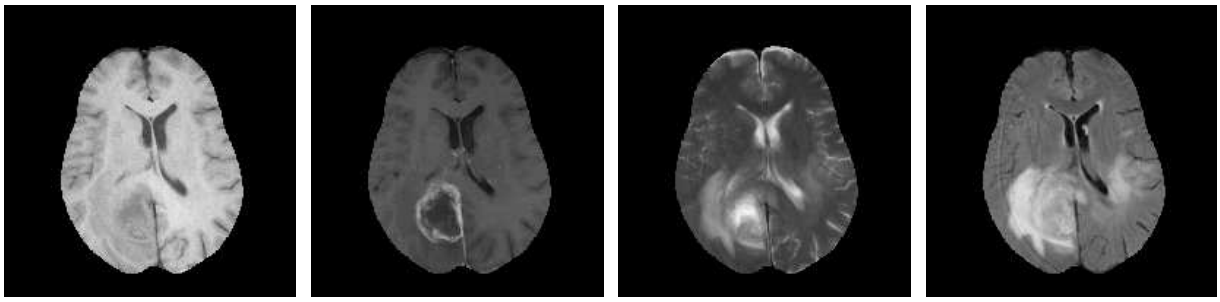

Figure 3.2: Example of an axial slice of different MRI sequences showing a GBM tumour. From left to right: T_1, T_{1c}, T_2 and FLAIR image.

3.2 BRAin Tumour Segmentation (BRATS) 2013 dataset

In order to make the results of this M.Sc. Thesis comparable and to provide an overview of the performance of unsupervised classification techniques for brain tumour segmentation, we used the public multi-modal BRAin Tumour Segmentation (BRATS) 2013 dataset. This dataset was released for the international NCI-MICCAI 2013 Grand Challenges in Image Segmentation of MICCAI Conference[a].

The BRATS 2013 dataset provides two sets for the development of automated brain tumour segmentation methods. The training set consists of multi-contrast MR scans of 30 real glioma patients: 20 with High Grade (HG) glioma tumour and 10 with Low Grade (LG) glioma tumour. Additionally, 25 synthetic cases of HG and LG glioma tumour are provided. For each patient of the training dataset, both real and synthetic, expert manual segmentations are provided. The test set consists of multi-contrast MR scans of 10 HG glioma patients without the expert labellings. An evaluation web page was published the day of the Challenge to upload and assess the quality of the test segmentations.

Table 3.1 summarizes the distribution of cases provided in the BRATS 2013 dataset.

	Total	Real		Synthetic	
		HG	LG	HG	LG
Training	80	20	10	25	25
Test	10	10			

Table 3.1: Distribution of the number of patients provided in the BRATS 2013 dataset.

For each patient, T_1-weighted, T_2-weighted, contrast-enhanced T_1-weighted (T_{1c}) and FLAIR MR images were provided. All images were linearly co-registered to the T_{1c} sequence, skull stripped, and interpolated to 1 mm isotropic resolution. No inter-patient registration was made to put all the images in a common reference space.

BRATS 2013 manual expert annotations comprise five intensity levels:

Class 1: Non-brain, non-tumour, necrosis, cyst and haemorrhage. Further on red color.

Class 2: Surrounding edema. Further on green color.

Class 3: Non-enhancing tumour. Further on blue color.

[a] `https://www.virtualskeleton.ch/BRATS/Start2013`

Class 4: Enhancing tumour core. Further on yellow color.

Class 0: Everything else.

Figure 3.3 shows an example of a patient from the training set of the BRATS 2013 dataset. The same slice is showed in different MRI sequences also with the manual expert labelling proposed by the radiologists.

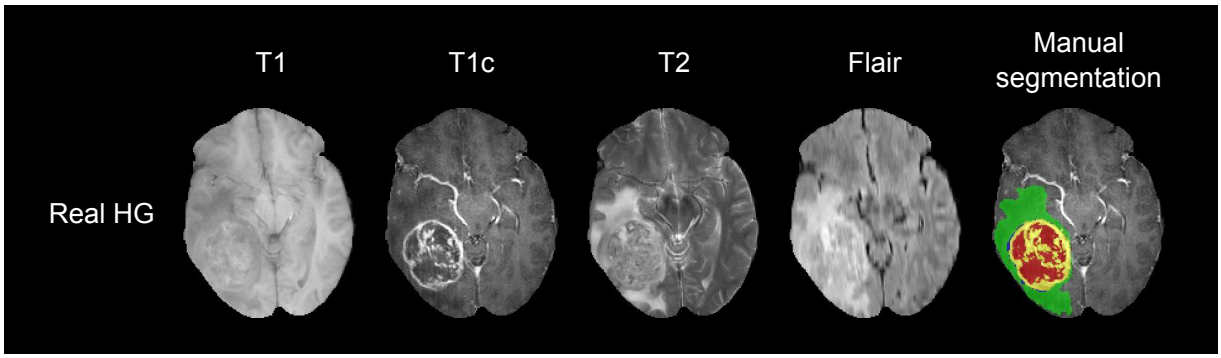

Figure 3.3: Example of an axial slice of different MRI sequences from a patient of the BRATS 2013 dataset, showing a GBM tumour and the manual segmentation provided by expert radiologist. From left to right: T_1, T_{1c}, T_2, FLAIR, and manual segmentation overlaid on the T_{1c} image.

Due to the unsupervised condition of the method proposed in this M.Sc. Thesis, no training set is required to learn the models of the classes. Hence, we only used the test partition of the BRATS 2013 dataset to develop the proposed unsupervised segmentation approach and to evaluate it in the same conditions than the supervised methods. Thus, we provide an assessment of the performance of different unsupervised segmentation algorithms in this public real dataset, and a comparison of these techniques with the state-of-the-art supervised segmentation methods that participated in the challenge.

3.3 International Consortium for Brain Mapping (ICBM) templates

Under the International Consortium for Brain Mapping (ICBM) Project, unbiased standard MR templates of normal brain volumes were provided by the McConnell Brain Imaging Centre in 2009[b].

These templates comprise the average of 152 healthy brains, non-linearly registered, bias field corrected and at different resolutions and symmetry conditions. T_1-weighted, T_2-weighted and Proton Density MR modalities were provided with their probability maps for the White Matter (WM), Gray Matter (GM) and Cerebro-Spinal Fluid (CSF) tissues. 6 different templates regarding to the resolution, preprocessing and symmetry conditions of the hemispheres are available:

[b] `http://www.bic.mni.mcgill.ca/ServicesAtlases/HomePage`

ICBM 2009a Nonlinear Symmetric: 1x1x1 mm^3 symmetric template with tissue probability maps. T_2 relaxometry, lobe atlas and different brain masks were also provided.

ICBM 2009a Nonlinear Asymmetric: 1x1x1 mm^3 asymmetric template with tissue probability maps. T_2 relaxometry, lobe atlas and different brain masks were also provided.

ICBM 2009b Nonlinear Symmetric: 0.5x0.5x0.5 mm^3 symmetric template with tissue probability maps.

ICBM 2009b Nonlinear Asymmetric: 0.5x0.5x0.5 mm^3 asymmetric template with tissue probability maps.

ICBM 2009c Nonlinear Symmetric: 1x1x1 mm^3 symmetric template with tissue probability maps. T_2 relaxometry, lobe atlas and different brain masks were also provided. Sampling is different from 2009a template.

ICBM 2009c Nonlinear Asymmetric: 1x1x1 mm^3 asymmetric template with tissue probability maps. T_2 relaxometry, lobe atlas and different brain masks were also provided. Sampling is different from 2009a template.

In this M.Sc. Thesis, we used the ICBM 2009c template for the postprocessing stage. Figure 3.4 shows an overview of the different atlases provided with the ICBM 2009c template.

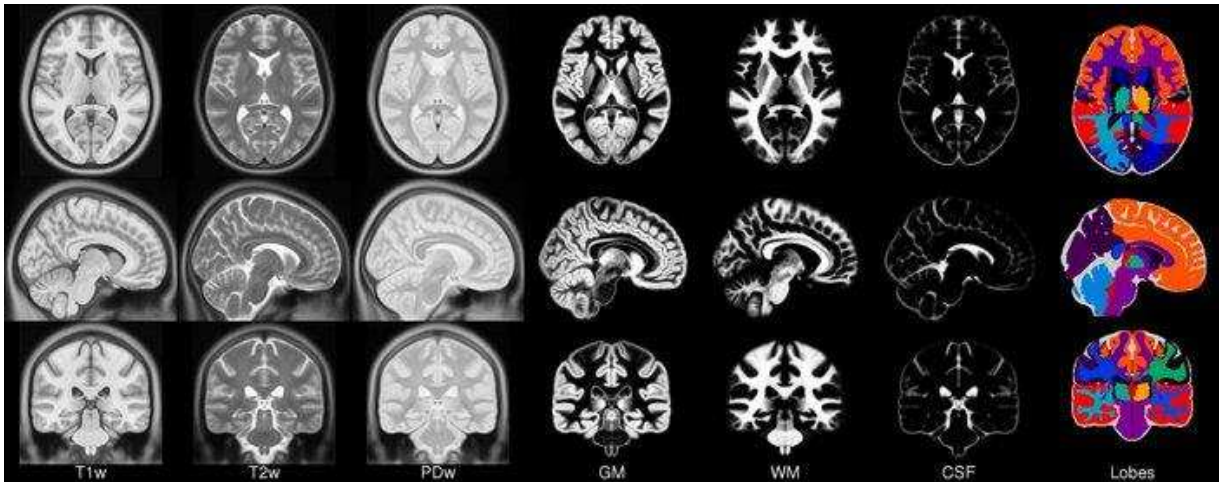

Figure 3.4: ICBM 2009c template. Top row shows the axial views of the different atlases provided with the template. Middle row shows the sagital view of the atlases, while bottom row shows the coronal view. From the left column to the right column: T_1-weighted sequence; T_2-weighted sequence; Proton Density sequence, GM tissue probability map; WM tissue probability map, CSF tissue probability map; Lobes segmentation

Tissue probability maps indicates for each voxel v of the brain the probability to belong to a normal tissue $T = \{WM, GM, CSF\}$, thus

$$\sum_{t \in T} p(t|v) = 1$$

Chapter 4

Methods

This chapter introduces the methods used in the M.Sc. Thesis to design the unsupervised segmentation approach for GBM tumour. We begin with a section about the MRI preprocessing techniques used to correct common artefacts of MR acquisitions and to enhance the information contained in the images. Then, a feature extraction and dimensionality reduction section is presented including the techniques used to extract discriminative features from the images and to reduce the high dimensionality of the data in order to simplify the inference of the segmentation algorithms. Then, the unsupervised classification algorithms are presented, separated by its structured or non-structured nature. All unsupervised techniques are postulated in terms of variants of generative mixture models and the EM algorithm. Then, the proposed method to automatically isolate the pathological classes related to the GBM tumour is presented to improve the biological interpretability of the results. Finally, the last section presents the evaluation strategies and figures of merit used to assess the performance of the method.

Figure 4.1 shows the graphical scheme of the proposed unsupervised GBM segmentation approach presented in this M.Sc. Thesis.

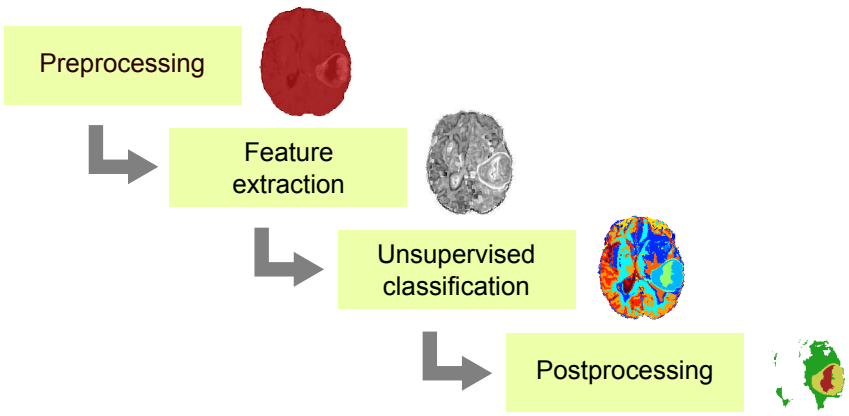

Figure 4.1: Proposed unsupervised GBM tissue segmentation pipeline.

4.1 MRI preprocessing

The first important stage in the GBM segmentation method is the preprocessing. MRI preprocessing is an active field of research that attempts to enhance and correct MR images for its posterior analysis. In an unsupervised segmentation approach this step becomes more important due to the absence of prior knowledge to guide the learning or segmentation process. Hence, common artefacts such as noise or inhomogeneities may rise as erroneous classes increasing the importance of an effective MRI preprocessing. We propose the following scheme for preprocessing MR images before the segmentation of the different tissues in the brain: 1) Denosing; 2) Registration; 3) Skull-stripping; 4) Bias field correction; 5) Super resolution.

There is no standard preprocessing pipeline accepted for MRI data, however several reasons can be adduced to justify the proposed preprocessing scheme. In order to avoid the propagation of noise to posterior stages of the preprocessing, the denoising step is first carried out. Next, the registration is performed to put all MR images in a common reference space, which allows the computation of a unique skull stripping mask, valid for all the MRI acquisitions of the patient. Hence, prioritizing the registration to the skull stripping step, the computational cost of the preprocessing is reduced. Bias field correction also benefits from the previous registration step when temporal sequences such as Perfusion Weighted Images (PWI) are handled. In these cases, the bias field correction must be adapted to not consider each dynamic of the sequence independently, and hence not destroying the temporal information. Finally, the super resolution usually requires that the different MRI acquisitions are registered in a common voxel space and the images are as much filtered as possible to take advantage of the self similarity between all the acquisitions of a patient. Furthermore, the proposed preprocessing pipeline is also valid for both MR pathological and non-pathological images, and for supervised and unsupervised segmentation approaches.

4.1.1 Denoising

Denoising is a standard preprocessing task for MRI manipulation, which aims to reduce or ideally remove the noise from an image. Although MRI noise has been usually modelled as a Gaussian distribution, by definition MRI noise follows a Rician distribution [31]. Diaz *et al.* [32] presented in 2011 a comprehensive analysis of different denoising methods, discussing their weaknesses and strengths. Figure 4.2 shows an example of MRI denoising. Top row shows a noisy original T_1 sequence. Middle row shows the denoised version of the T_1 sequence. Bottom row shows the residuals corresponding to the noise of the image.

A common drawback of denoising methods concerns the removal of high frequency signal components of the images during the filtering. Recent approaches such as the Non Local Means (NLM) introduced by Buades *et al.* [33] has improved the existing techniques for MR data. Based on this approach, Manjón *et al.* [34] introduced a variant of the filter, which does not assume an uniform distribution of the noise over the image, thereby adapting the strength of the filter depending on a local estimation of the noise. The filter also deals with both correlated Gaussian and Rician noise. In this M.Sc. Thesis, the approach proposed by Manjón *et al.* is used to remove the noise of the MR images.

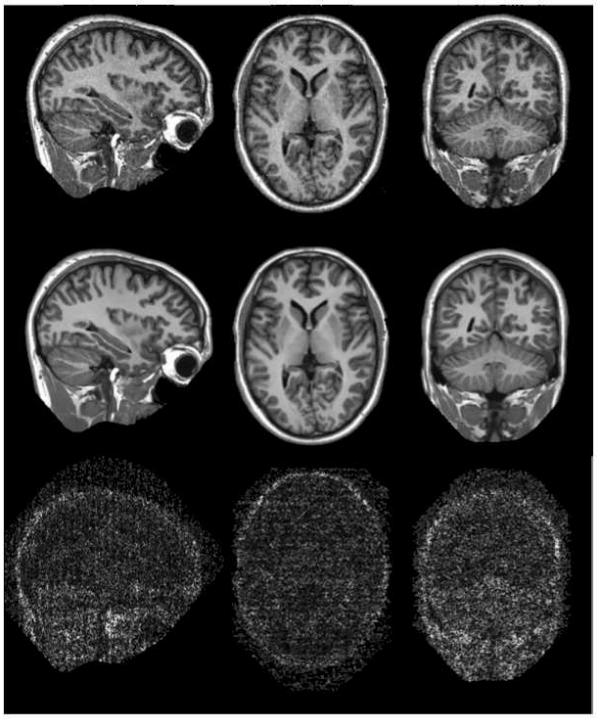

Figure 4.2: Example of denosing of a T_1 sequence. Top row: noisy T_1; Middle row: Denoised T_1; Bottom row: Residuals obtained after the denoising process.

4.1.2 Registration

Registration is another common process in medical imaging field. The standard protocol for GBM diagnosis concerns the acquisition of several MR images. When a multi-modal analysis is performed, it is mandatory that all MR images are in a common reference space [7]. In other words, it is required that the same area of the brain is represented by the same voxel positions in all MR sequences, to avoid introducing inconsistencies or mixtures of tissues from different MRI acquisitions. In normal clinical practice, MR images are not usually registered, so to ensure the voxel correspondences the registration step needs to be carried out. The T_1 sequence of the patient is usually used as a reference to register the rest of MR images. In brain tumour lesions, affine or linear registration methods are preferred to not deform the areas of the tumour, thereby keeping intact the lesion.

An extensive evaluation of 14 non-linear registration methods was carried out in 2009 by Klein *et al.* [35]. This work concluded that SyN algorithm [36] implemented in the Advanced Normalization Tools (ANTS) suite was one of the best registration algorithms in terms of accuracy, flexibility and efficiency. In this sense, we propose the use of ANTS to perform the registration of the different MRI acquisition for the posterior segmentation. Figure 4.3 shows an example of the registration process.

Although our preprocessing scheme includes the registration step, the BRATS dataset comes with an intra-patient MRI registration. The evaluation web-page provided to assess the performance of the proposed segmentations also holds the ground truth of each BRATS patient registered to its corresponding MR images. Hence, no registration should be done in order to not deform the images and its segmentations, thereby ensuring a correct

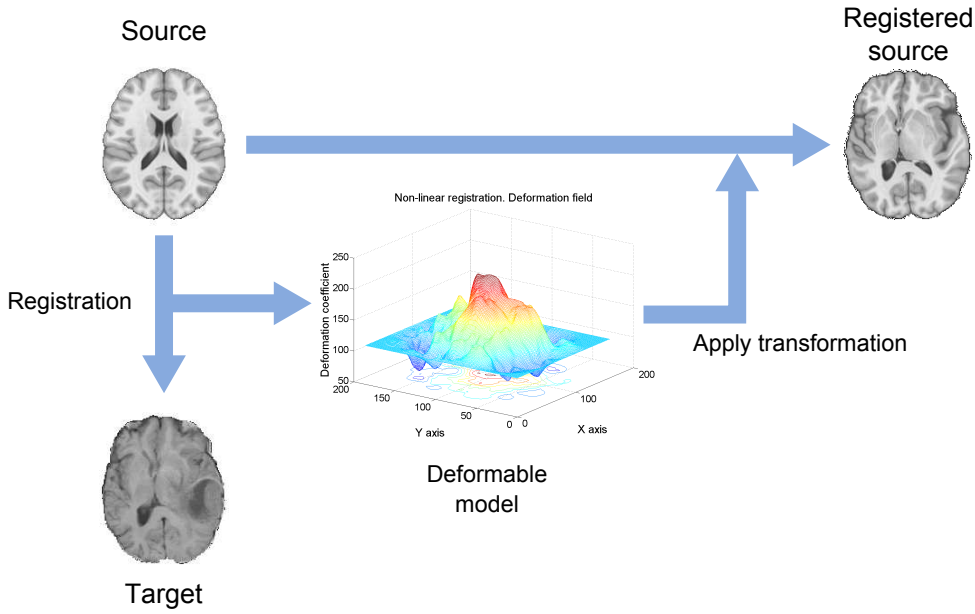

Figure 4.3: Example of a non-linear MRI registration process.

evaluation.

4.1.3 Skull stripping

Skull stripping process comprises the removal of skull, extra-meningial and non-brain tissues from the MRI acquisitions. In [37], a review of different skull-stripping methods was presented, however it did not include recent methods such as the ones provided in the Brain Suite Software[a] [38], or the RObust Brain EXtraction (ROBEX) method proposed by Iglesias *et al.* [39], which claims to provide significantly improved performance in a multi-dataset evaluation, against six popular skull stripping methods. Figure 4.4 shows an example of the intra-cranial mask computed through the skull stripping method provided in the Brain Suite Software.

MR images of the BRATS dataset are also skull stripped, however we detected several cases with partial inclusion of areas of the cranium that should be removed. For that purpose, the Brain Suite Sofware was used to automatically compute a skull stripping mask over the T_1 sequences, and removed the non desired cranium areas.

4.1.4 Bias field correction

Intensity inhomogeneity is another common artefact present in MRI acquisitions. Magnetic field inhomogeneities are an unavoidable effect in MRI, which generates a low frequency signal that corrupts the images affecting their intensity levels. Hence, the same tissue in the brain could present different gray level distributions across the image, introducing inconsistencies that directly affects the segmentation methods. Typically, automated segmentation approaches are based on the assumption that the brain tissues present the same distribution of intensity among the image. Therefore, a preprocessing step is

[a] `http://brainsuite.org/`

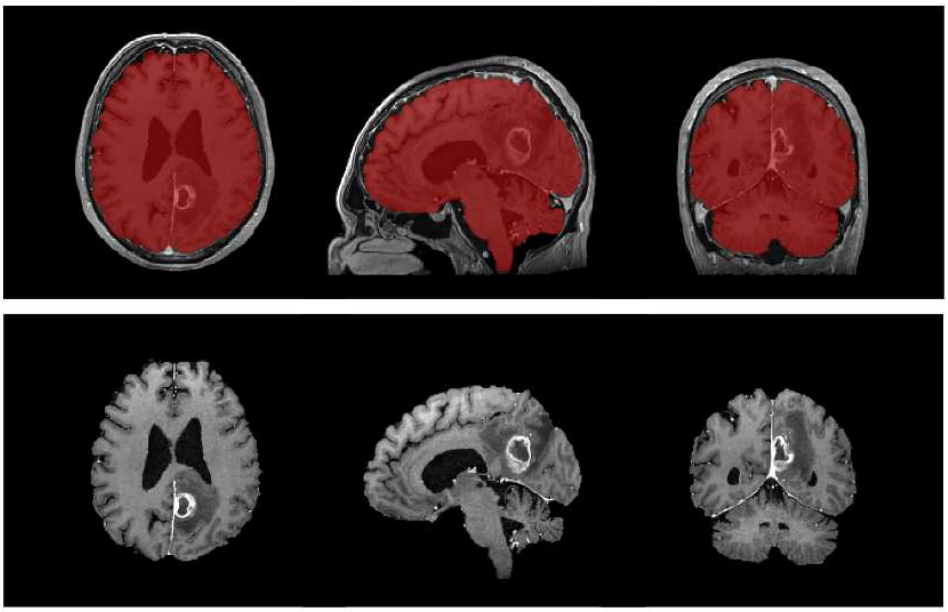

Figure 4.4: Example of a skull stripping process on a high resolution 3D T_{1c} sequence. Mask computed with the skull stripping method provided with the Brain Suite Software.

needed to correct the bias field before the segmentation. Figure 4.5 shows an example of a T_{1c} MRI sequence presenting magnetic field inhomogeneities and the corrected sequence with its estimation of the bias of the magnetic field.

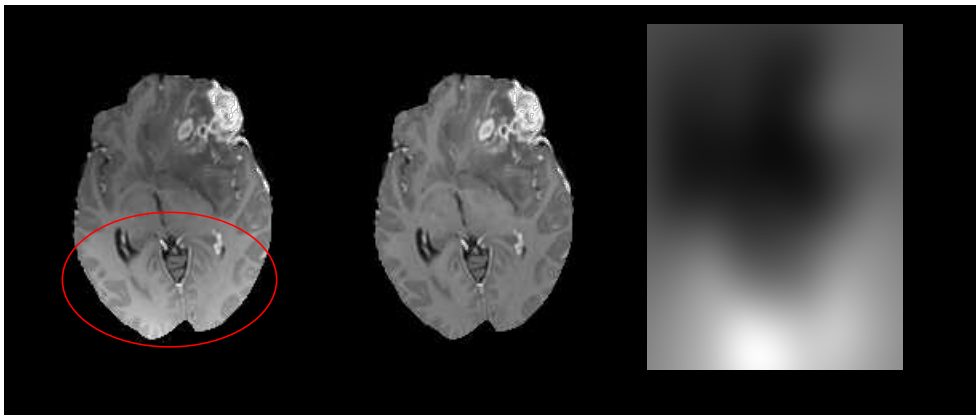

Figure 4.5: Example of bias field correction on a T_{1c} sequence. From left to right: Original T_{1c} sequence with magnetic field inhomogeneities artefact; T_{1c} sequence after the bias field correction; Estimation of the bias of the magnetic field produced by the MRI machine.

The popular non-parametric non-uniform intensity normalization N3 algorithm was proposed in 1998 by Sled *et al.* [40], becoming a reference technique for bias field correcting because of no tissue model was needed to perform the correction. Tustison *et al.* [41] proposed in 2010 a new implementation of N3 called N4, which improves the N3 algorithm with a better B-spline fitting function and a hierarchical optimization scheme for the bias field correction. N4 is used in this M.Sc. Thesis for MRI magnetic field inhomogeneity correction.

4.1.5 Super resolution

In a brain tumor lesion protocol, several MR sequences are commonly acquired at different resolutions, thereby introducing spatial limitations when a multi-modal study is performed. In these cases, an upsampling or interpolation is needed to set a common voxel space for all images. Classical interpolations, such as linear, cubic or splines interpolation could rise as a solution for the problem, but at the cost of introducing common artefacts in the images such as partial volume effects or stair-case artefacts. In contrast, more powerful and sophisticated methods such as super resolution could improve classical interpolation methods by reconstructing the low resolution images, recovering its high frequency components. Several super resolution schemes for MR imaging are available in the literature [42, 43, 44, 45].

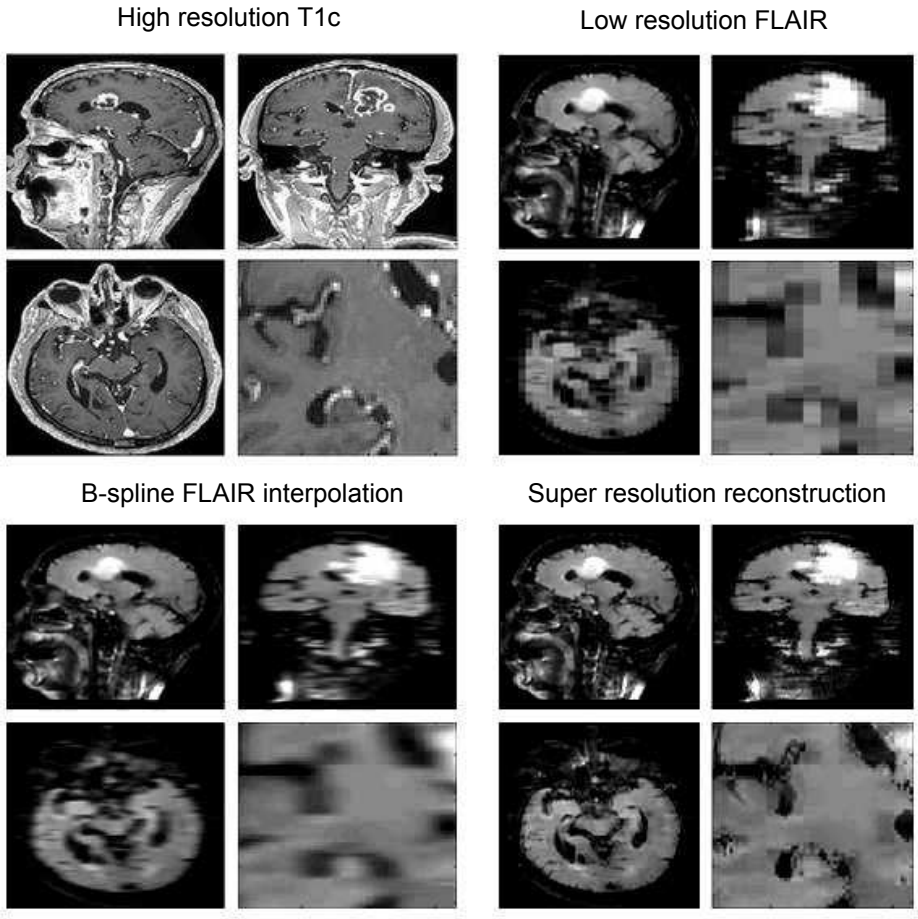

Figure 4.6: Example of super resolution reconstruction of a low resolution FLAIR sequence using information of a high resolution T_{1c} sequence from the same patient. Comparison between classic b-spline interpolation of the low resolution FLAIR sequence and the super resolution reconstruction.

BRATS dataset comes with a $1mm^3$ isotropic voxel size resolution achieved through classic interpolation. Such interpolation could be improved through an iterative super resolution reconstruction process to recover the high frequency components of the image. Specifically, we use the super resolution algorithm proposed by Manjón *et al.* [46], which

exploits the the self-similarity present in MR images through a patch-based non-local reconstruction process.

4.2 Feature Extraction and Dimensionality Reduction

Feature extraction comprises the process of obtaining new features from the MR images to improve the discrimination between different tissues in the posterior segmentation. Although MRI intensities are the most common features used to discriminate between tissues in the brain, it has been shown that including texture features in combination with MR intensities increases the performance of brain tumour segmentation [47, 48]. Several approaches to extract textures from images are proposed in the past years. Robert Haralick provided in [49] the reference paper for analysis of textures in images. Later, Van Gool *et al.* [50] also reviewed the algorithms for texture analysis and both authors agreed in classifying the texture extraction methods in two categories: Statistical methods and structural methods. Structural methods are better suited to regular large patterns of texture, however statistical texture features present better performance for non-regular micro textures in images. In this M.Sc. Thesis we adopted the statistical texture feature analysis for texture representation.

Several approaches has been proposed to extract statistical texture features from images [51, 52]. We implemented the first order texture features, often called histogram derived metrics or first central moments, as a method for texture representation. Such features consist on the computation of the histogram in a local 3D neighbourhood centred at each voxel of the MR image, and then compute the mean, variance, skewness and kurtosis of such histogram. Thus, we computed the texture features for all the MRI sequences of a patient (T_1, T_{1c}, T_2 and FLAIR), using a local 3D neighbourhood of $5 \times 5 \times 5$ voxels for all the non-background voxels of the images. Figure 4.7 shows an example of the first order texture features computation on a T_{1c} sequence.

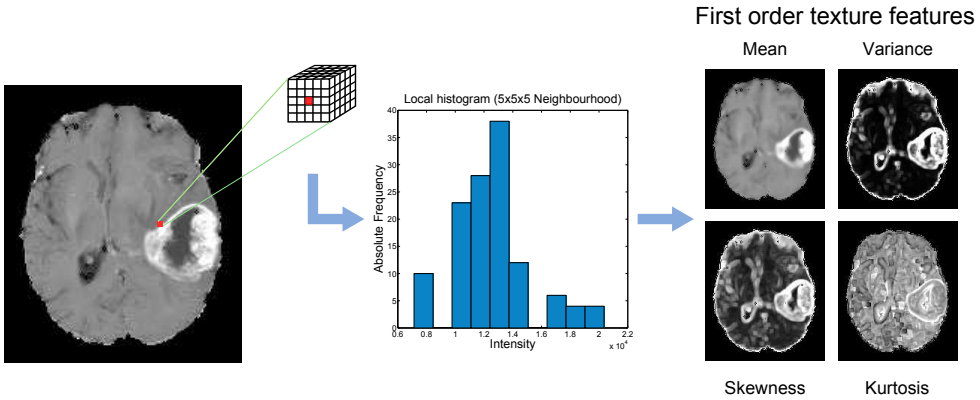

Figure 4.7: Example of first order texture features computation on a T_{1c} sequence of a patient of the BRATS dataset.

Besides the texture feature images, an additional image named T_{1d} is computed. The T_{1d} image is obtained from the absolute difference of the T_{1c} and T_1 sequence.

$$T_{1d} = \mid T_{1c} - T_1 \mid$$

This image highlights the contrast enhanced areas of the T_{1c} image, such as the active tumour, and also helps in the discrimination of WM and GM tissues. First order texture features are also extracted from the T_{1d} image in the same manner as for the other MRI acquisitions.

As a result, a set $\mathbf{I}$ of 25 images (3D volumes) are obtained from each patient.

$$\mathbf{I} = (T_1, T_{1c}, T_2, FLAIR, T_{1d}, \mu_{T_1}, ..., \sigma_{T_1}, ..., \gamma_{T_1}, ..., \kappa_{T_{1d}}) \ / \ \mathbf{I} \in \mathbb{R}^{X \times Y \times Z \times D}$$

where μ, σ, γ, and κ prefixes refers to the mean, variance, skewness and kurtosis texture features respectively, X, Y, Z are the dimensions of the images (equal for all images after the registration), and D refers to the dimensions of each voxel, id est (that is) (i.e) the number of different images or features ($D = 25$).

Dimensionality reduction is the process of efficiently represent the original high dimensional data into a lower dimensional space, but retaining or increasing its most relevant information. Several dimensionality reduction algorithms have been presented in the past years, differentiating between linear and non-linear approaches. Figure 4.8 shows a taxonomy of the most popular dimensionality reduction algorithms.

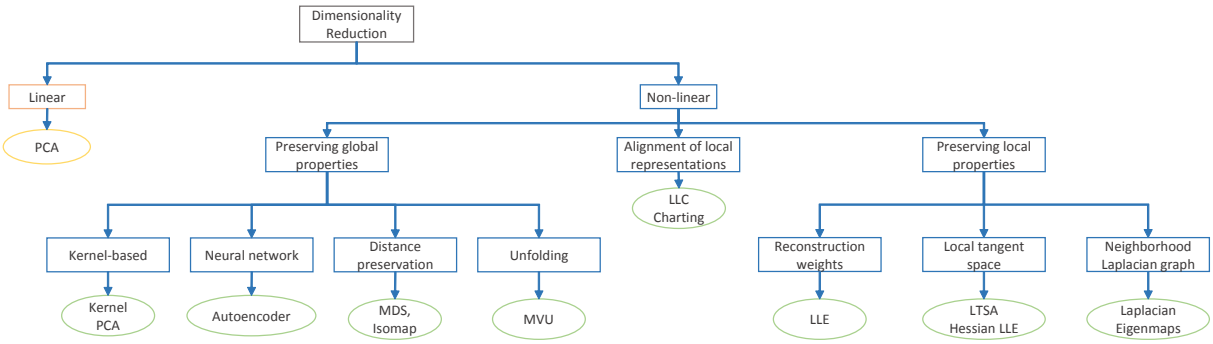

Figure 4.8: Taxonomy of dimensionality reduction algorithms.

In this M.Sc. Thesis, PCA is used to reduce the voxel dimensionality of the input data $\mathbf{I}$. PCA is a linear reduction technique, which seeks for an orthogonal transformation of a number of hypothetical correlated variables into a smaller number of uncorrelated variables called *principal components*. Such principal components are sorted in terms of amount of variance explained from the data, becoming the first component the one that accounts for as much variability as possible. Projecting the original data over the first components of PCA (number of components less than the number of original dimensions), a dimensionality reduction is achieved by preserving the most variability in the data as possible.

As PCA is based on the variance of the data to perform the reduction, we must exclude the variance texture feature of the set $\mathbf{I}$ to avoid that first components of PCA retain only this feature. Thus, the dimensions are directly reduced to $D = 20$ and the $\mathbf{I}$ set is then formed by:

$$\mathbf{I} = (T_1, T_{1c}, T_2, FLAIR, T_{1d}, \mu_{T_1}, ..., \gamma_{T_1}, ..., \kappa_{T_{1d}}) \ / \ \mathbf{I} \in \mathbb{R}^{X \times Y \times Z \times D}$$

PCA is finally applied to retain the 99% of the variance of the data, achieving a reduction $D = 5$ for all patients of the BRATS 2013 dataset. An slice example of the feature extraction and PCA dimensionality reduction of a patient is shown in Figure 4.9.

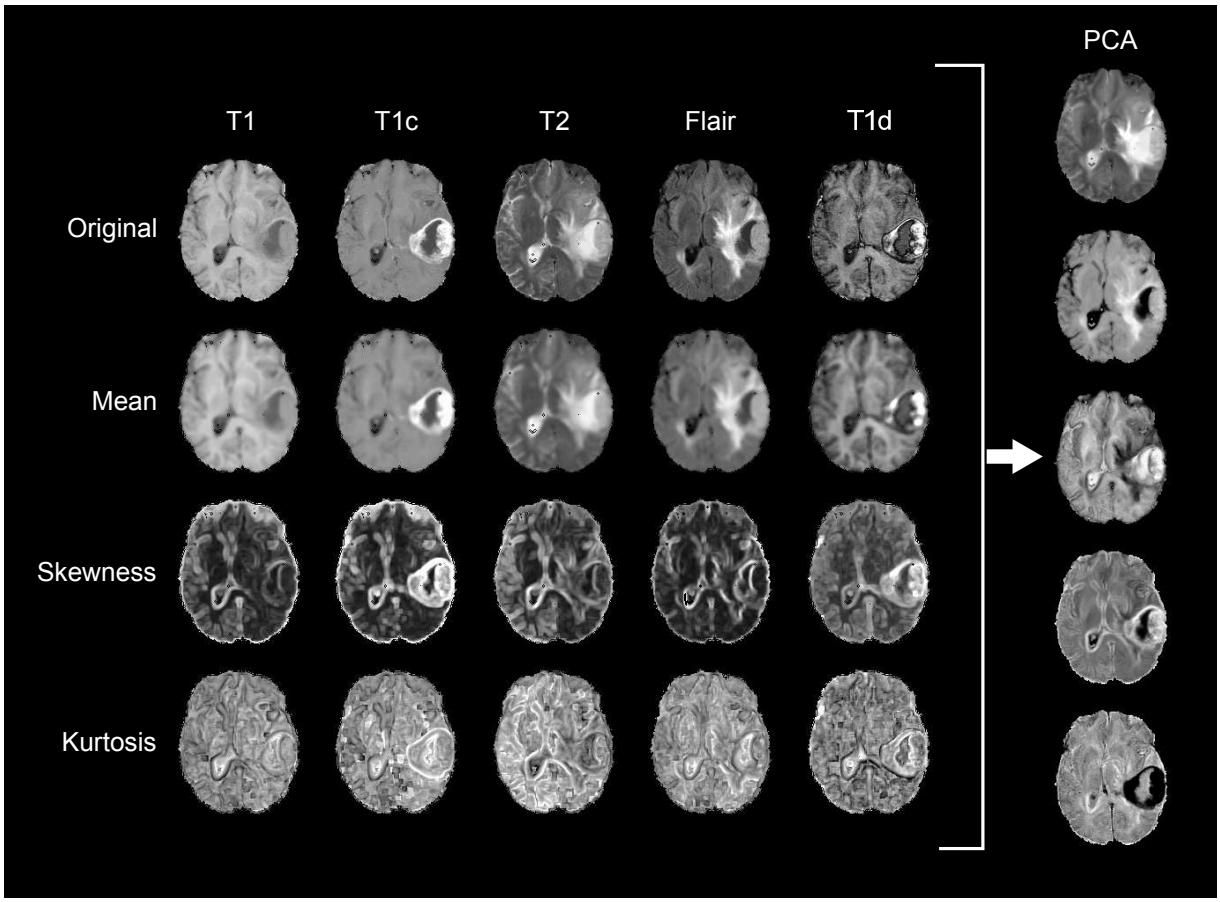

Figure 4.9: Example of feature extraction and dimensionality reduction of a patient of the BRATS dataset.

4.3 Unsupervised voxel classification

Unsupervised learning is a branch of Machine Learning (ML) which calls for finding a hidden structure in the input data, often called clusters, formed by natural groupings of observations. The major difference between supervised and unsupervised learning relies in that the second receives an additional sequence of desired outputs from which to learn the relations with the inputs and the discriminant models to distinguish them. Conversely, unsupervised learning does not require a set of labelled data. Instead, unsupervised learning examines the input data in order to find clusters of homogeneous information that represents each one a different class.

We evaluated the most popular unsupervised classification algorithms to segment both normal and pathological tissues in the brain. We divided the algorithms in two groups: Structured and non-structured classification algorithms. Non-structured algorithms classify data assuming an iid condition between the observations (voxels) of the dataset. Structured prediction covers the range of algorithms that assume and model data with a specific structure, such as an image, i.e assume conditional dependency between the observations. Under the non-structured paradigm, we evaluated three methods: K-means, Fuzzy K-means and Gaussian Mixture Model (GMM) clustering. In the structured prediction case we evaluated Gaussian Hidden Markov Random Field (GHMRF) as the archetype of unsupervised structured learning.

BRATS 2013 dataset comprises 5 classes to be segmented, which in some cases a single class encloses several types of tissues (for example 0 class). This intra-class heterogeneity severely affects the performance and interpretation of the unsupervised results. While supervised approaches can address this heterogeneity by enforcing the algorithm to learn a model that represents the data enclosed in a class, unsupervised approach often solves the problem by separating the heterogeneous classes in different clusters. Thus, the unsupervised approach often requires to estimate more than the initially defined classes. However, it is then mandatory to design a postprocessing stage after the initial segmentation, to improve the interpretability of the results, and merge the possible redundant classes. Often, prior knowledge about the task could give an orientation of the number of classes that should be estimated. In this case we assume that at least 7 different tissues exist in the brain, which are:

Tissue 1: Class 1 of BRATS 2013 dataset. Non-brain, non-tumour, necrosis, cyst and haemorrhage.

Tissue 2: Class 2 of BRATS 2013 dataset. Surrounding edema.

Tissue 3: Class 3 of BRATS 2013 dataset. Non-enhancing tumour.

Tissue 4: Class 4 of BRATS 2013 dataset. Enhancing tumour.

Tissue 5: White Matter (WM).

Tissue 6: Gray Matter (GM).

Tissue 7: Cerebro-Spinal Fluid (CSF).

Moreover, due to the intra-class heterogeneity presented in some classes, we assume that each class is modelled through a mixture of 2 Gaussians, giving a total of 14 classes. Therefore, 14 classes are estimated for each unsupervised classification algorithm.

A well-known requirement of unsupervised learning algorithms is the good initial seeding. Although global minima is not usually reached even if a good initialization is provided, a bad initialization can lead the model to a hard local minimum, thereby providing a poor segmentation. Several strategies such as multiple replications or intelligent initial seeding are proposed to palliate this effect. In this M.Sc. Thesis, we implemented the *K-means++* algorithm proposed in [53], which provides an initialization that attempts to avoid local minima. Additionally, we generated 100 different *K-means++* initializations, and run each unsupervised segmentation algorithms with each initialization. Finally, the best solution for each algorithm is chosen attending to the following criterion:

- Lowest intra-cluster sums of point-to-centroid distances is used for K-means and Fuzzy K-means algorithms.
- Lowest Negative Log-Likelihood value is used for GMM clustering and GHMRF.

4.3.1 Expectation-Maximization (EM) algorithm

Expectation-Maximization (EM) [54, 55] is an algorithm proposed by Arthur Dempster, Nan Laird and Donald Rubin in 1977, which is used to find the maximum likelihood parameters of a statistical model in cases where latent variables and unknown parameters are involved. In our case, EM is used to estimate probabilistic generative mixture models where both labels and parameters of the underlying model are unknown.

Let $X = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N)$ the set of observations to be classified, where $\mathbf{x}_n \in \mathbb{R}^D$ represents a feature vector of D dimensions for observation n. Let $Y = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_N)$, where $\mathbf{y}_n \in \{0,1\}^{|C|}$ an unit-length multinomial variable that indicates the class for the observation n. Let $\mathcal{C} = \{1, ...C\}$ the set of all possible classes.

From the statistical point of view, the classification problem is often addressed through the Bayes decision rule, which ensures that the minimization of the global risk is obtained by maximizing the posterior probability of the classes given the observations.

$$\hat{Y} = \arg\max_{Y} p_\theta(Y|X)$$

where θ represent the parameters of the underlying probabilistic model.

The supervised classification paradigm assumes that both observations X and labels Y are known and then, the *learning* process consists of estimating the parameters of the model θ that best fit the observed data. Thus, it is called that the supervised approach has the *complete dataset*, denoted by $\{X, Y\}$. Maximum likelihood is usually employed to estimate the parameters of the model, which considering the complete dataset becomes straightforward.

$$\mathcal{L}(\theta|X, Y) = p_\theta(X, Y)$$

When an unsupervised learning approach is adopted, both labels Y and parameters θ are not observed, so maximum likelihood estimation of the model is not directly possible.

Hence, both θ and Y are considered as latent or hidden variables, thus becoming the model

$$p_\theta(X) = \sum_Y p_\theta(X, Y)$$

The log-likelihood function is then given by

$$\mathcal{L}(\theta|X) = \log p_\theta(X) = \log \left\{ \sum_Y p_\theta(X, Y) \right\}$$

which becomes in a more complex solution.

In such cases, EM algorithm can be used to iteratively obtain the most likelihood parameters of the model and the non-observed values of the labels Y. EM consist of two steps: the Expectation (E) step, in which the expected value of the joint distribution of the observations and the labels is computed given the parameters of the model and the posterior probability at the current iteration, and the Maximization (M) step, in which an updating of the old parameters of the model is performed based on the recently computed posterior probability in the E step.

Thus, the general form of the EM algorithm is as follows

Initialization step: Initialize $\theta^{(0)}$ parameters.

Expectation step:

$$\begin{aligned} Q(\theta|\theta^{(k)}) &= E_{p_{\theta^{(k)}}(Y|X)} \left(\log p_\theta(X, Y) \right) \\ &= \sum_Y p_{\theta^{(k)}}(Y|X) \left(\log p_\theta(Y) + \log p_\theta(X|Y) \right) \end{aligned}$$

Maximization step:

$$\theta^{(k+1)} = \arg\max_\theta Q(\theta|\theta^{(k)})$$

Convergence step: Stop if $\mathcal{L}(\theta^{(\mathbf{k+1})}|X) - \mathcal{L}(\theta^{(\mathbf{k})}|X) \leq \epsilon$; otherwise $k = k + 1$ and go to **Expectation step**.

In practice, the $Q(\theta|\theta^{(k)})$ function is not usually computed in the Expectation step. Instead, only the posterior probability (which is the unknown distribution at this point) is required to estimate the new parameters of the model in the Maximization step. Hence, in real implementations of the algorithm, only the posterior probability is computed in the Expectation step.

EM algorithm is not guaranteed to converge to a global maxima of the likelihood function, so several heuristics or strategies has been proposed such as multiple replications, intelligent initial seeding or simulated annealing.

4.3.2 Non-structured Gaussian mixture model

A non-structured mixture model refers to a probabilistic model where the observations are assumed to be independent and identically distributed (iid). In this cases, the probabilistic model has the form

$$p_\theta(X, Y) \overset{iid}{=} \prod_n p_\theta(\mathbf{x}_n, \mathbf{y}_n)$$

Following a generative process, $p(\mathbf{y}_n)$ and $p(\mathbf{x}_n|\mathbf{y}_n)$ are generated as

$$p_{\mathbf{p}}(\mathbf{y}_n) = \prod_{c \in C} p_c^{y_{nc}} \text{ governed by } \mathbf{p} = (p_1, ..., p_C)$$

$$p_{\theta'}(\mathbf{x}_n|\mathbf{y}_n) = \prod_{c \in C} p_{\theta'}(\mathbf{x}_n|c)^{y_{nc}} \text{ governed by } \theta'$$

Thus, the parameters that governs the probabilistic model are $\theta = (\mathbf{p}, \theta')$.

In the unsupervised classification paradigm where both parameters θ and labels Y are not observed, the model is expressed as

$$\begin{aligned}
p_\theta(X) &= \sum_Y p_\theta(X, Y) \\
&\overset{iid}{=} \sum_{\mathbf{y}_1} \cdots \sum_{\mathbf{y}_N} \prod_n p_\theta(\mathbf{x}_n, \mathbf{y}_n) \\
&= \left[\sum_{\mathbf{y}_1} p_\theta(\mathbf{x}_1, \mathbf{y}_1)\right] \left[\sum_{\mathbf{y}_2} \cdots \sum_{\mathbf{y}_N} \prod_{n=2} p_\theta(\mathbf{x}_n, \mathbf{y}_n)\right] \\
&= \cdots \\
&= \prod_n \sum_{\mathbf{y}_n} p_\theta(\mathbf{x}_n, \mathbf{y}_n)
\end{aligned}$$

A typical assumption is to model the data as a mixture of (independent) multivariate Gaussians. In this case, the θ parameters are defined as

$$\theta = (\mathbf{p}, \theta')$$

$$\mathbf{p} = (p_1, ..., p_C)$$

$$\theta' = (\theta'_1, ..., \theta'_C) = (\mu_1, ..., \mu_C; \Sigma_1, ..., \Sigma_C)$$

The EM algorithm for non-structured Gaussian mixture models performs as follows

Initialization step: Initialize $\theta^{(0)}$ parameters (i.e each p_c, μ_c and Σ_c).

Expectation step:

$$\begin{aligned}
Q(\theta|\theta^{(k)}) &= E_{p_{\theta^{(k)}}(Y|X)} \left[\log p_\theta(X, Y)\right] \\
&\overset{iid}{=} \sum_n E_{p_{\theta^{(k)}}(\mathbf{y}_n|\mathbf{x}_n)} \left[\log p_\theta(\mathbf{x}_n, \mathbf{y}_n)\right] \\
&= \sum_n \sum_c y_{nc}^{(k)} \left(\log p_c + \log p_{\theta'}(\mathbf{x}_n|c)\right)
\end{aligned}$$

where $y_{nc}^{(k)}$ is the posterior probability of $\mathbf{x}_n$ being generated from component c

$$y_{nc}^{(k)} = E_{p_{\theta^{(k)}}(\mathbf{y}_n|\mathbf{x}_n)}[y_{nc}] = p_{\theta^{(k)}}(y_{nc} = 1|\mathbf{x}_n) = \frac{p_c^{(k)} p_{\theta'}^{(k)}(\mathbf{x}_n|c)}{\sum_{c'} p_{c'}^{(k)} p_{\theta'}^{(k)}(\mathbf{x}_n|c')}$$

Maximization step:

$$\begin{aligned}\theta^{(k+1)} &= \arg\max_{\theta} Q(\theta|\theta^{(k)}) \text{ , subject to } \sum_c p_c = 1 \\ &= \arg\max_{\theta} \max_{\lambda} Q(\theta|\theta^{(k)}) - \lambda\left(\sum_c p_c - 1\right)\end{aligned}$$

where taking derivatives of Q w.r.t $\mathbf{p}$, θ' and λ and equating them to zero gives

$$\begin{aligned}p_c^{(k+1)} &= \frac{1}{N}\sum_n y_{nc}^{(k)} \\ \mu_c^{(k+1)} &= \frac{1}{\sum_n y_{nc}^{(k)}}\sum_n y_{nc}^{(k)}\ \mathbf{x}_n \\ \Sigma_c^{(k+1)} &= \frac{1}{\sum_n y_{nc}^{(k)}}\sum_n y_{nc}^{(k)}\left(\mathbf{x}_n - \mu_c^{(k+1)}\right)\left(\mathbf{x}_n - \mu_c^{(k+1)}\right)^t\end{aligned}$$

Convergence step: Stop if $\mathcal{L}(\theta^{(\mathbf{k+1})}|X) - \mathcal{L}(\theta^{(\mathbf{k})}|X) \leq \epsilon$; otherwise $k = k + 1$ and go to **Expectation step**.

4.3.2.1 K-means

K-means [56, 57] is an unsupervised non-structured iterative partitional clustering based on a distance minimization criterion. Its aim is to divide the data space into C clusters $(C \leq N)$ $J = \{J_1, J_2, ..., J_C\}$ so as to each observation belongs to the cluster with nearest centroid. The distance criterion minimized by K-means is

$$\arg\min_{J} \sum_{c}^{C} \sum_{\mathbf{x}_n \in J_c} \parallel \mathbf{x}_n - \mu_c \parallel^2$$

From a statistical point of view, the K-means approach is equivalent to find the most likelihood parameters of a mixture of multivariate Gaussians [58] (each Gaussian represents a class), assuming a shared identity covariance matrix and uniform prior probabilities for all Gaussians. Thus, each class follows $\theta' \sim \mathcal{N}(\mu_{\mathbf{c}}, I)$ and prior probability for each class is $p_c = 1/\mid \mathcal{C} \mid$. Moreover, the iterative approach followed by K-means is also demonstrated a special limit of the EM algorithm, called *Hard-EM*. In this variant of the algorithm, the observations are assigned *hardly* to the closest Gaussian of the mixture, assuming a posterior probability of $p_\theta(y_{nc} = 1|\mathbf{x}_n) = 1$, instead of computing its real probability.

Thus, the Hard-EM version implemented by K-means performs as follows

1. **Initialize** parameters $\theta^{(0)}$ (i.e initialize $\mu_{\mathbf{c}}$ for each class c)

2. **Hard E step:** Given the current parameters $\theta^{(k)}$ at iteration k, compute the following Q function:

$$Q(\theta \mid \theta^{(k)}) = \sum_n \sum_c y_{nc}^{(k)} \left(\log p_c + \log p_{\theta'}(\mathbf{x_n}|c)\right)$$

where:

$$y_{nc}^{(k)} = \begin{cases} 1, & \text{if } \| \mathbf{x_n} - \mu_{\mathbf{c}}^{(\mathbf{k})} \|^2 < \min_{c' \neq c} \| \mathbf{x_n} - \mu_{\mathbf{c'}}^{(\mathbf{k})} \|^2 \\ 0, & \text{otherwise} \end{cases}$$

3. **M step:** Make a guess $\theta^{(k+1)}$ by choosing θ that maximizes the Q function:

$$\theta^{(k+1)} = \arg\max_{\theta} Q(\theta|\theta^{(k)}) \text{ subject to } \sum_c p_c = 1$$

where the parameter update is computed as:

$$\mu_c^{(k+1)} = \frac{1}{\sum_n y_{nc}^{(k)}} \sum_n y_{nc}^{(k)} \mathbf{x}_n$$

4. **Stop** if samples do not change class; **otherwise** go to step 2.

4.3.2.2 Fuzzy K-means clustering

Like K-means, Fuzzy K-means [59, 60] also proposes a mixture of multivariate Gaussian distribution assuming a shared identity covariance matrix and uniform prior probabilities for all classes. However, Fuzzy K-means differs from K-means in which the assignment of a observation to a class is not *hard* but *fuzzy*. This means that each observation now keeps a degree of membership to each Gaussian (related to its posterior probability) instead of a unique assignment with posterior probability of 1. In the same manner as K-means, the aim is to divide the data space into C clusters ($C \leq N$) $J = \{J_1, J_2, ..., J_C\}$, but it also provides a vector $\mathbf{u}_n$ for each observation, which determines the membership degree of the observation n to the different clusters. The new distance minimization criterion followed by Fuzzy K-means is

$$\arg\min_J \sum_c^K \sum_{\mathbf{x}_n \in J_c} u_{nc}^m \| \mathbf{x}_n - \mu_c \|^2 \qquad 1 \leq m < \infty$$

where m controls the degree of fuzziness of the cluster c, typically set to 2 in absence of domain knowledge, and u_{nc} is defined as

$$u_{nc} = \frac{1}{\sum_j^K \left(\frac{\|\mathbf{x}_n - \mu_c\|^2}{\|\mathbf{x}_n - \mu_j\|^2} \right)^{\frac{2}{m-1}}}$$

where u_{nc} is proportional to the posterior probability of cluster c given the observation n, $u_{nc} \sim p_\theta(y_{nc}|\mathbf{x}_n)$.

The EM version proposed by Fuzzy K-means is then formulated as follows

1. **Initialize** parameters $\theta^{(0)}$ (i.e initialize $\mu_{\mathbf{c}}$ for each class c)

2. **Hard E step:** Given the current parameters $\theta^{(k)}$ at iteration k, compute the following Q function:

   $$Q(\theta \mid \theta^{(k)}) = \sum_n \sum_c y_{nc}^{(k)} \left(\log p_c + \log p_{\theta'}(\mathbf{x_n}|c)\right)$$

   where:

   $$y_{nc}^{(k)} \equiv u_{nc}$$

3. **M step:** Make a guess $\theta^{(k+1)}$ by choosing θ that maximizes the Q function:

   $$\theta^{(k+1)} = \arg\max_{\theta} Q(\theta|\theta^{(k)}) \text{ subject to } \sum_c p_c = 1$$

   where the parameter update is computed as:

   $$\mu_c^{(k+1)} = \frac{1}{\sum_n u_{nc}^{(k)}} \sum_n u_{nc}^{(k)} \, \mathbf{x}_n$$

4. **Stop** if samples do not change class; **otherwise** go to step 2.

4.3.2.3 GMM clustering

Gaussian Mixture Model (GMM) clustering is the generalization of K-means and Fuzzy K-means algorithms. The hard constraints imposed by K-means and Fuzzy K-means related to the prior probabilities and covariance matrices derives in linear decision boundaries of the data space, which often lead to weak classifiers with a low performance. Also the hard assignment of the observations to the classes may derive in noisy classifications instead of soft decision boundaries. Thus, a natural improvement is to consider free covariance matrices for each class, non equal prior probabilities and soft assignment based on posterior probabilities for each class. This approach builds a more flexible model, able to represent non-linear decision boundaries.

Attending to these new conditions, the EM algorithm is then the proposed in the subsection 4.3.2, which performs as follows

1. **Initialize** parameters $\theta^{(0)}$ (i.e initialize $\mu_{\mathbf{c}}$, Σ_c and p_c for each class c)

2. **Hard E step:** Given the current parameters $\theta^{(k)}$ at iteration k, compute the following Q function:

   $$Q(\theta \mid \theta^{(k)}) = \sum_n \sum_c y_{nc}^{(k)} \left(\log p_c + \log p_{\theta'}(\mathbf{x_n}|c)\right)$$

   where:

   $$y_{nc}^{(k)} = \frac{p_c^{(k)} p_{\theta'}^{(k)}(\mathbf{x}_n|c)}{\sum_{c'} p_{c'}^{(k)} p_{\theta'}^{(k)}(\mathbf{x}_n|c')}$$

3. **M step:** Make a guess $\theta^{(k+1)}$ by choosing θ that maximizes the Q function:

$$\theta^{(k+1)} = \arg\max_{\theta} Q(\theta|\theta^{(k)}) \text{ subject to } \sum_c p_c = 1$$

where the parameter update is computed as:

$$p_c^{(k+1)} = \frac{1}{N}\sum_n y_{nc}^{(k)}$$

$$\mu_c^{(k+1)} = \frac{1}{\sum_n y_{nc}^{(k)}}\sum_n y_{nc}^{(k)}\ \mathbf{x}_n$$

$$\Sigma_c^{(k+1)} = \frac{1}{\sum_n y_{nc}^{(k)}}\sum_n y_{nc}^{(k)}\left(\mathbf{x}_n - \mu_c^{(k+1)}\right)\left(\mathbf{x}_n - \mu_c^{(k+1)}\right)^t$$

4. **Stop** if $\mathcal{L}(\theta^{(\mathbf{k+1})}|X) - \mathcal{L}(\theta^{(\mathbf{k})}|X) \leq \epsilon$; **otherwise** go to step 2.

4.3.3 Structured Gaussian mixture model

In a structured mixture model, no iid assumption between observations is made, so conditional dependencies between the observations are considered. Thus, the mixture model should handle the set of labels and observations jointly. Regarding that both parameters (θ) and labels (Y) are not observed, the mixture model is expressed as

$$p_\theta(X) = \sum_Y p_\theta(X,Y) = \sum_Y p(Y)p_\theta(X|Y)$$

In order to introduce structured dependencies between the observations of the model, Markov Random Fields (MRFs) are usually used. MRFs are probabilistic undirected graphical models, in which conditional dependencies between random variables are explicitly denoted via a undirected and cyclic graph. The vertices of the graph represent the random variables of the model, and the edges of the graph represent the conditional dependencies among these variables. Figure 4.10 shows an example of a undirected graphical model, representing a generative probabilistic model of the form

$$p_\theta(X,Y) = p_(Y)\prod_n p_{\theta'}(x_n|y_n)$$

From a generative point of view, the MRF is defined over the prior probability $p(Y)$, which according to the Hammersey-Clifford theorem [61], follows a Gibbs distribution of the form

$$p(Y) = \frac{1}{Z}\exp\left(-U(Y)\right)$$

where $U(Y)$ is called the energy function and Z is the partition function that ensures the distribution to sum 1.

$$Z = \sum_{Y'}\exp\left(-U(Y')\right)$$

Hammersey-Clifford theorem also states that $p(Y)$ can be factorized over the cliques of the graphical model. A clique is defined as a subset of vertices in the graph, such that

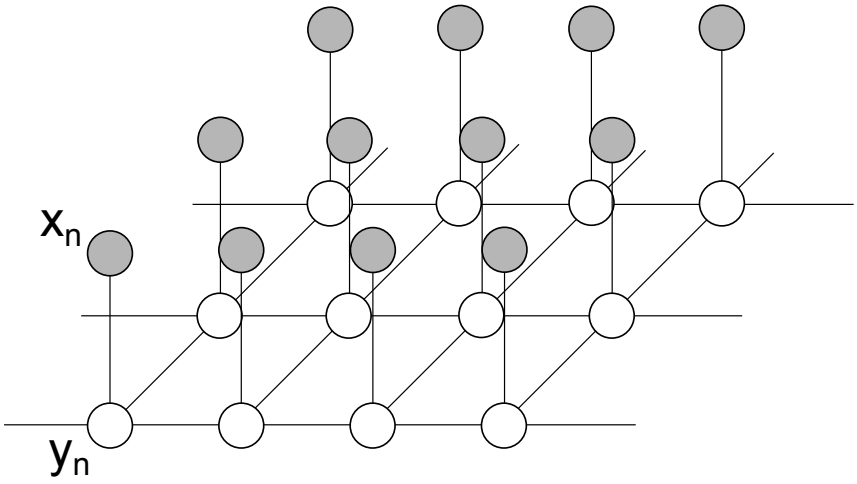

Figure 4.10: Undirected graphical model representing a generative probabilistic model. Lattice formed between white vertices determines the conditional dependencies between the random variables of the model and is modelled through a Markov Random Field (MRF). Gray vertices represent the observations of the model and edges that that connect gray nodes with white nodes represent the conditional dependency between observations and the random variables.

there exists a link between all pairs of vertices in the subset, i.e is fully connected. A random variable of the model is then considered independent given its cliques.

Let Q the set of all cliques of the graph. The energy function $U(Y)$ is then defined as

$$U(Y) = \sum_{q \in Q} \Psi_q(Y)$$

Nowadays, if complexity is considered, the inference algorithms for MRFs can only do an optimization job for undirected graphs with cliques of order 2, (pairwise cliques). Hence, the most commonly used graphical model is the *Ising model*. The Ising model defines a graph lattice where conditional dependencies of each variable are expressed in terms of its orthogonal adjacent neighbourhood. Figure 4.10 shows an Ising model for a 2D plane, represented by the graph defined by the white vertices.

The clique factorization for the *Ising model* performs as follows

$$U(Y) = \sum_{q \in Q} \Psi_q(y_n, y_m) = \sum_{q \in Q} \beta_q \delta(y_n, y_m)$$

where $y_n, y_m \in q$, $\Psi_q(y_n, y_m)$ is the clique potential for clique q, β_q is a weight defined for such clique (in our case always $\beta = 1$) and δ is a function that measures the dissimilarity between classes of the variables involved in the clique, typically

$$\delta(y_n, y_m) = \begin{cases} 0, & \text{if } y_n = y_m \\ 1, & \text{otherwise} \end{cases}$$

Although MRF strictly refers to the prior probability of the generative model, typically, the class conditional probability is also expressed in terms of energy functions, so $p_\theta(X \mid Y)$ is usually rewritten as

$$p_\theta(X|Y) = \frac{1}{Z} \exp\left(-U_\theta\left(X \mid Y\right)\right)$$

where $U_\theta\left(X \mid Y\right)$ is proportional to the class conditional probability of the observations given the current parameters and labelling, and Z is again a partition function to ensure the distribution to sum 1. As the prior probability $p(Y)$ determines the conditional dependencies between the observations of the model, the class conditional probability $p(X|Y)$ could be assumed iid between the observations, thus

$$U_\theta(X|Y) = \prod_n U_\theta(\mathbf{x}_n|\mathbf{y}_n)$$

As a result, the structured mixture model is therefore defined as

$$p_\theta(X) = \sum_Y \frac{1}{Z} \exp\left(-U_\theta\left(X \mid Y\right) - U\left(Y\right)\right)$$

Exact inference on this model is intractable due to the sum over all possible configuration of labels denoted by Z, which is a $\#P-complete$ problem. However, approximate efficient algorithms are available to compute the best labelling when pairwise conditional dependencies are considered. Thus, $p_\theta(X)$ is approximated by the maximum

$$p_\theta(X) \approx \max_Y \frac{1}{Z} \exp\left(-U_\theta\left(X \mid Y\right) - U\left(Y\right)\right)$$

Although these, inference algorithms do not compute $p_\theta(X)$, indeed they provide the best labelling $\hat{Y}$ and its final energy value

$$\begin{aligned}\hat{Y} &= \arg\max_Y(-U_\theta(X|Y) - U(Y)) \\ &= \arg\min_Y(U_\theta(X|Y) + U(Y))\end{aligned}$$

where following an unit-length multinomial notation, $\hat{Y} \in \{0,1\}^{(N\times C)}$, i.e, an indicator matrix that specifies the class c for each observation n.

Several algorithms are proposed for the inference of MRFs such as Iterated Conditional Modes (ICM), Monte Carlo Sampling or Graph cuts. In this M.Sc. Thesis we used the algorithm proposed by Komodakis *et al.* [62, 63], based on a combination of Graph cuts with primal-dual strategies.

4.3.3.1 GHMRF

GHMRF is the Gaussian unsupervised variant of MRF. The term HMRF refers to a hidden generative structured model based on a MRF prior, where labels are not observed. The Gaussian assumption of the class conditional probabilities of the model finally coins it as GHMRF. Likewise GMM clustering, GHMRF considers free covariance matrices for each class and non equal prior probabilities. However, as complete inference of the model is not possible, only a Hard EM version is available to estimate the parameters of the structured model. Thus, the EM version proposed for the GHMRF is

1. **Initialize** parameters $\theta^{(0)}$ (i.e initialize $\mu_\mathbf{c}$, Σ_c and p_c for each class c)

2. **Hard E step:** Given the current parameters $\theta^{(k)}$ at iteration k, compute the following Q function:

$$Q(\theta \mid \theta^{(k)}) = \sum_n \sum_c \hat{Y}_{nc}^{(k)} \left(\log p_c + \log p_{\theta'}(\mathbf{x_n}|c)\right)$$

where:

$$\hat{Y}^{(k)} = \arg\min_Y \left(U_\theta(X|Y) + U(Y)\right)$$

3. **M step:** Make a guess $\theta^{(k+1)}$ by choosing θ that maximizes the Q function:

$$\theta^{(k+1)} = \arg\max_\theta Q(\theta|\theta^{(k)}) \text{ subject to } \sum_c p_c = 1$$

$$p_c^{(k+1)} = \frac{1}{N} \sum_n y_{nc}^{(k)}$$
$$\mu_c^{(k+1)} = \frac{1}{\sum_n y_{nc}^{(k)}} \sum_n y_{nc}^{(k)}\ \mathbf{x}_n$$
$$\Sigma_c^{(k+1)} = \frac{1}{\sum_n y_{nc}^{(k)}} \sum_n y_{nc}^{(k)} \left(\mathbf{x}_n - \mu_c^{(k+1)}\right) \left(\mathbf{x}_n - \mu_c^{(k+1)}\right)^t$$

4. **Stop** if samples do not change class; **otherwise** go to step 2.

4.4 Automatic tumour classes isolation

Unlike supervised learning, unsupervised segmentation produces a partitioning of the data space into several classes, but each class without *semantic* sense. In other words, in the unsupervised approach, class labels between different segmentations may not always represent the same tissue, complicating its biological interpretation. Hence, tumour classes isolation is mandatory to provide a powerful and competitive unsupervised brain tumour segmentation method. We propose the following postprocessing pipeline to automatically isolate pathological classes:

1. Identify WM, GM and CSF classes
2. Remove outlier classes
3. Mixture classes by statistical distribution similarities

4.4.1 Identify WM, GM and CSF classes

In order to identify WM, GM and CSF classes, and to isolate the pathological classes in the brain, the tissue probability maps provided by the ICBM 2009c template are used (see section 3.3). However, ICBM template represents a healthy brain. Hence tissue probability maps do not consider any other tissue rather than normal tissues $T = \{WM, GM, CSF\}$.

Therefore, it is required to correct these tissue probability maps by introducing an additional tissue denoted by L, to deal with the lesion area. Consequently, WM, GM and CSF probability maps should be renormalized by spreading its probability in the lesion area on a new map, such that

$$T' = \{WM, GM, CSF, L\}$$

$$\sum_{t \in T'} p(t|v) = 1$$

To correct the tissue probability maps, we first performed a non-linear registration of the ICBM T_1 sequence to the patient T_1 sequence and then applied the non-linear transformation obtained through the registration to the ICBM tissue probability maps. We employ cross-correlation metric with the SyN algorithm [36] implemented in the ANTS suite to perform the registration. Next, a roughly approximate mask of the lesion area is computed. The delineation performed by the expert radiologist of the margins of the tumour is usually based on the hyper-intensity areas in the T_2 and T_{1c} sequences [4]. Following a similar criterion, we compute an approximate mask of the lesion area by retrieving the histograms of the FLAIR and T_{1c} sequences and selecting those voxels with an intensity level higher than the median plus the standard deviation of any histograms. Next, holes of each 2D axial plane of the current computed mask are filled and voxels that fall in the perimeter of the volume are automatically removed. Finally, the lesion mask is used to set an ϵ value in each normal tissue probability maps in the area defined by the mask. It is worth noting that this mask do not delimits or restricts the shape of the classes provided by the unsupervised segmentation, but serves to identify the pathological area, and then which classes of the segmentation primarily explains the lesion. Figure 4.11 shows the computation process of the corrected tissue probability maps.

Based on the corrected tissue probability maps, we identify which classes of a given segmentation mainly explain the normal tissues T. For a given segmentation S, a normal tissue $t \in T$ and for each class $c \in S$, we computed the following probability:

$$p(c|t, S) = \frac{\sum_{v,\ S(v)=c} p(t|v)}{\sum_v p(t|v)}$$

where v denotes a voxel of the volume.

The $p(c|t, S)$ determines the probability for each class c to explain the normal tissue t. Hence, in order to remove the classes that primarily represent the normal tissue t, we sort all classes in descending order by the $p(c|t, S)$

$$\mathcal{C}_t = \{c_i \ \mid \ p(c_i|t, S) \geq p(c_{i+1}|t, S),\ 1 \leq i < C\}$$
$$\mathcal{P}_t = \{p(c_i|t, S) \ \mid \ p(c_i|t, S) \geq p(c_{i+1}|t, S),\ 1 \leq i < C\}$$

Next, we compute the cumulative sum of $\mathcal{P}_t$

$$\Phi(i) = \sum_{j=1}^{i} \mathcal{P}_t(j), \qquad 1 \leq i < C$$

and delete the first classes until the 0.8 of the cumulative sum is reached.

$$\mathcal{Z}_t = \{\mathcal{C}_t(i+1) \ \mid \ \Phi(i) > 0.8,\ 1 \leq i < C\}$$

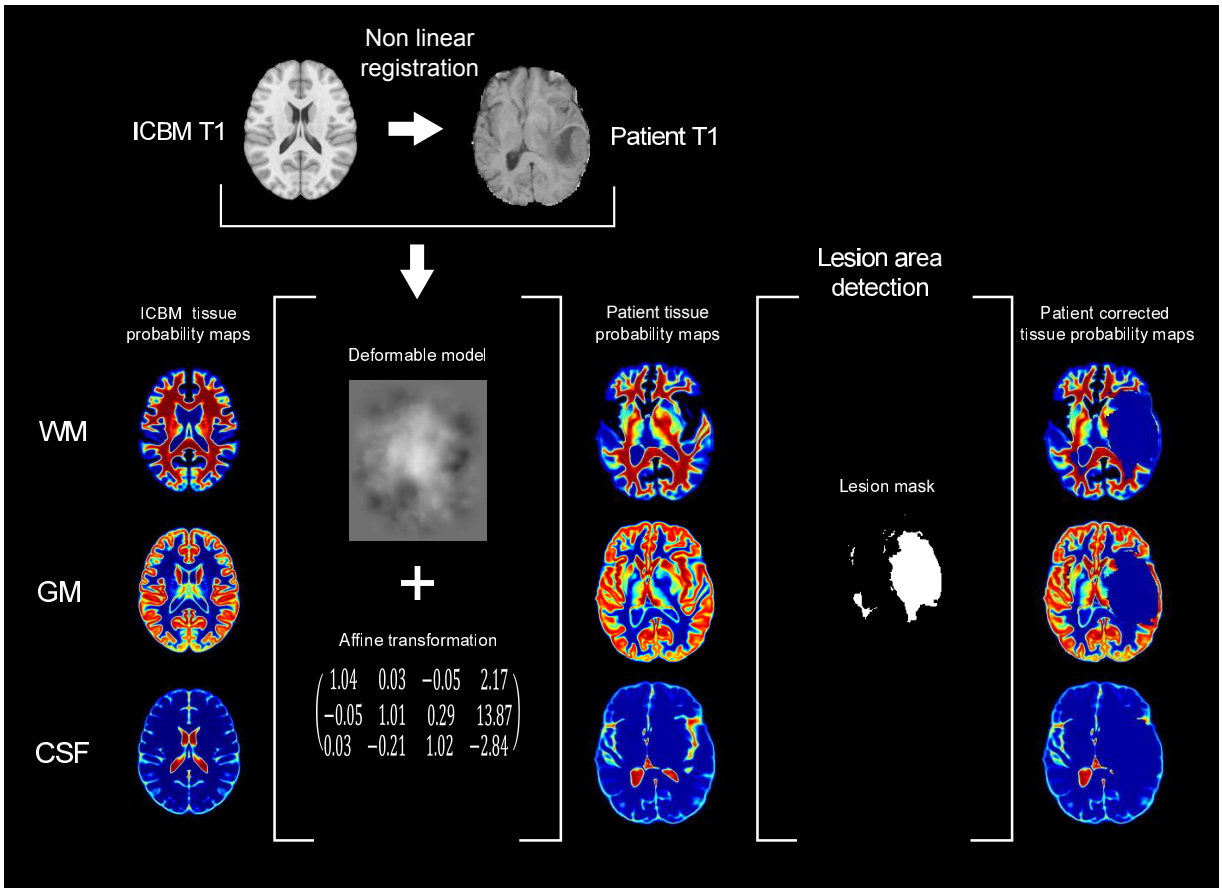

Figure 4.11: Patient specific tissue probability maps computation and subsequent lesion area correction.

We choose 0.8 as a reasonably value to explain each normal tissue t through a combination of classes. Note that a threshold of 1 implies deleting all the classes of the segmentation, due to each class always retains a minimum probability to belong to a normal tissue. Moreover, the tissue probability maps are obtained through a non-linear registration of a healthy template to a pathological brain, and a posterior correction of the tissue probability maps. Such process introduces unavoidable errors that should be considered when a threshold is defined to identify the non-pathological classes. Thus, in major cases, a 0.8 threshold provide a high confidence degree to identify the normal classes of the segmentation.

Finally, we repeated independently this procedure for each tissue $t \in \{WM, GM, CSF\}$ to isolate the pathological classes. The intersection of the sets obtained for each normal tissue removes the normal tissue classes, and provides the final pathological classes defined by the set $\mathcal{Z}$

$$\mathcal{Z} = \mathcal{Z}_{WM} \cap \mathcal{Z}_{GM} \cap \mathcal{Z}_{CSF}$$

4.4.2 Remove outlier classes

The process of identifying and removing the normal tissue classes (WM, GM and CSF) may leave some spurious classes that should be deleted. We find that these classes frequently appear in the perimeter of the brain or in a very low percentage of occurrence with respect to the rest of classes of the segmentation. The classes located at the perimeter of the brain usually represent the intensity gradient between the brain and the background or the partial volume effects that the super resolution cannot remove. The smaller classes often represent outlier voxels in terms of abnormal intensity values, usually produced by unavoidable artefacts in the MR acquisition.

In order to delete the perimeter unwanted classes, we first compute a binary mask of the perimeter of both hemispheres. Next, we dilate such mask in order to cover a wider area greater than one voxel. Finally, for each class after the 4.4.1 step, we compute its connected components and delete such connected components that falls into the perimeter mask with more than the 50% of its area.

In order to remove the smaller classes, we first compute the percentage of occurrence of each class over the whole segmentation and delete those ones with a percentage less than a 1%.

4.4.3 Mixture classes by statistical distribution similarities

The heterogeneity of the tumoral classes lead us to assume that each tissue of the brain is modelled through at least a mixture of two Gaussians. However, the unsupervised voxel classification provides a general mixture of Gaussians over the whole brain, that better fits the clusters of information present in the data. This means that a tissue may bind together more than two classes for its own representation, or by the opposite, requires an unique class if it is homogeneously enough. Thus, it is mandatory to provide a mechanism to find class similarities that allows a merging process that results in an homogeneous segmentation that correctly explains the final pathological tissues.

Based on the work proposed by Sáez *et al.* [64], we analysed the statistical distributions of the remaining classes after the 4.4.1 and 4.4.2 steps, to find possible mixtures of classes with similar distributions. We estimate a non-parametric probability density function for

each class through a kernel smoothing density estimation, and use the Jensen-Shannon divergence to measure its distances. Thus, we construct a pairwise matrix of statistical distribution distances between classes, and we use a Hierarchical Agglomerative Clustering (HAC) with an average link (Unweighted Pair Group Method with Arithmetic Mean (UPGMA)), to find similar classes.

Due to the BRATS 2013 labelling considers 4 pathological classes to be segmented, we enforce the clustering to return a maximum of 4 classes. Note that this is the maximum number of different classes that can be returned, however the method is able to return less than 4 classes if the HAC finds enough similarities to merge it. Moreover, this condition can be altered or removed if an exploratory approach is adopted, to find new sub-compartment segmentations.

Figure 4.12 shows an example of the full tumour classes isolation procedure, combining the normal tissue class removal (step 4.4.1), the outlier class removal (step 4.4.2) and the class merging by its statistical distribution similarities (step 4.4.3).

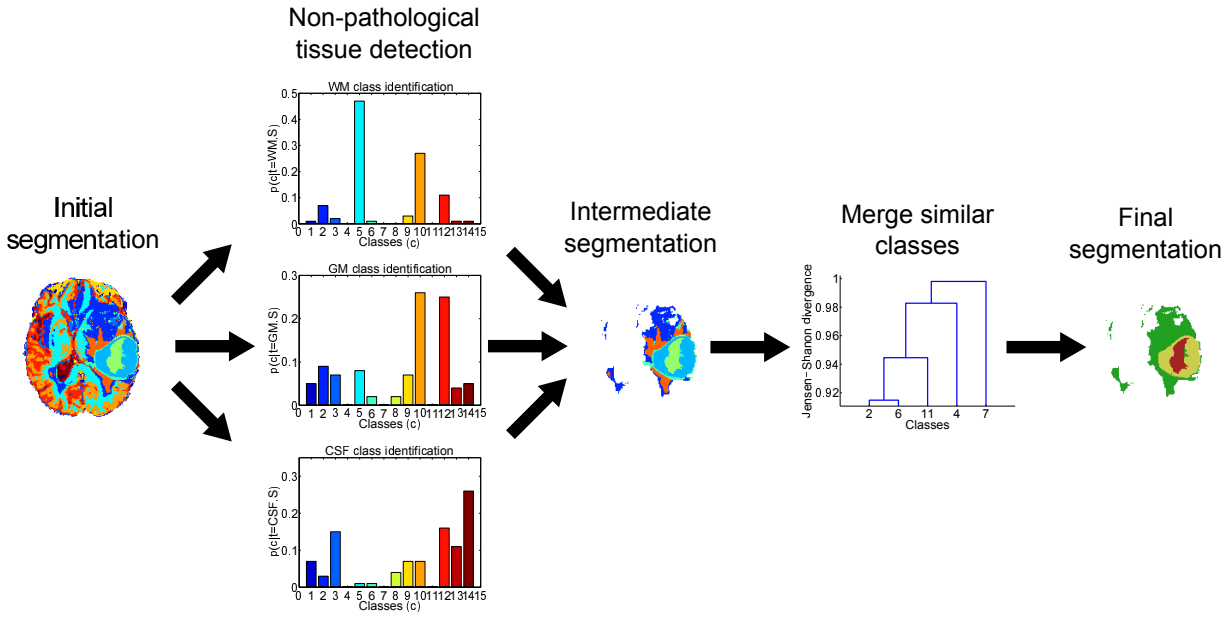

Figure 4.12: Automatic tumoral class isolation process.

Finally, we manually reorder the numbers of the classes to match the BRATS labelling. Note that this step is completely unavoidable due to the fact that the numbers of the classes becomes an arbitrary decision and can not be inferred by an unsupervised approach. Thus, this manual stage does not compromise the proposed automated methodology as it is a step only required for the concrete public dataset used to evaluate the method.

4.5 Evaluation

In this section, the evaluation strategies followed in this M.Sc. Thesis and proposed by the MICCAI Challenge to assess the quality of the segmentations are presented below. An evaluation web page[b] is provided for the organization committee of the Challenge to

[b] https://www.virtualskeleton.ch/BRATS/Evaluation2013

upload and evaluate the segmentations. Each segmentation should be identified by a code related to the FLAIR acquisition of the patient that originates the segmentation.

As this M.Sc. Thesis proposes an unsupervised segmentation method and in order to compare our results with the state-of-the-art supervised segmentation approaches, we only used the test set of the BRATS 2013 dataset. We directly segmented the test set with the four unsupervised segmentations algorithms (see section 4.3) in combination with the proposed preprocessing and postprocessing pipelines. The following subsections present the different subcompartments and metrics used to assess the quality of the segmentations.

4.5.1 Subcompartment evaluation

Three different sub-compartments are evaluated for the proposed segmentations. This evaluation strategy aims to provide a confident measure of the quality of the segmentation methods, avoiding global measures that could be influenced by some tissues predomination. The subcompartments are:

Complete tumour: Labels 1 + 2 + 3 + 4. Evaluation of the whole segmentation, including all the pathological tissues, i.e necrosis, cyst and haemorrhage (1), edema (2), non-enhancing tumour (3) and enhancing tumour (4).

Tumour core: Labels 1 + 3 + 4. Evaluation of tumoral tissues core. Includes necrosis, cyst and haemorrhage (1), non-enhancing tumour (3) and enhancing tumour (4).

Enhancing tumour: Label 4. Evaluation of only active/enhancing tumour (4).

Figure 4.13 shows an example of the subcomparment segmentations evaluated in the BRATS 2013 Challenge.

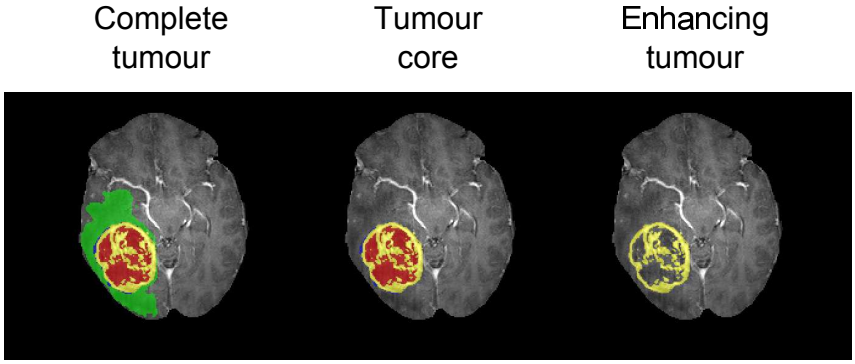

Figure 4.13: Example of different subcomparment segmentations evaluated in the BRATS 2013 Challenge.

4.5.2 Figures of merit

The figures of merit used to assess the quality of the segmentations are presented below. The S term refers to the proposed segmentation, while the GT term refers to the ground truth hold by the BRATS 2013 evaluation web page. Table 4.1 shows the confusion matrix of a binary classification problem and the performance metrics that can be calculated from it.

		True class		
		+	−	Row totals
Hypothesized class	$\hat{+}$	**T**rue **P**ositives	**F**alse **P**ositives	$\hat{\mathbf{P}}$
	$\hat{-}$	**F**alse **N**egatives	**T**rue **N**egatives	$\hat{\mathbf{N}}$
Column totals		**P**	**N**	

Table 4.1: Confusion matrix and performance metrics.

Sørensen-Dice coefficient: Similar than the Jaccard index, the Dice coefficient measures the set agreement between the proposed segmentation S and the ground truth GT. It computes the number of overlapped voxels between S and GT and divide it between the average of the sizes of S and GT. The Dice coefficient is the most typical measure to evaluate the quality of a segmentation. It ranges between 0 and 1, with 0 meaning absence of overlapping and 1 referring to complete perfect overlapping.

$$D = \frac{2 \mid S \cap GT \mid}{\mid S \mid + \mid GT \mid} = \frac{2(TP + TN)}{P + N + \hat{P} + \hat{N}}$$

Positive predictive value (PPV): Often called Precision, the PPV gives a metric to assess the susceptibility of the method to produce false positives in the segmentation. The statistic ranges from 0 to 1, with 0 meaning a low precision of the method and 1 a high precision, i.e, a low false positive rate.

$$PPV = \frac{TP}{TP + FP}$$

Sensitivity: Often called Recall, the sensitivity measures the success ratio of the method. Hence, the sensitivity computes the number of hits obtained by the method and divides it by the number of real positives of the ground truth. The measure ranges from 0 to 1, with 1 meaning a high success ratio.

$$S = \frac{TP}{TP + FN}$$

Cohen's Kappa index: A robust measure of the agreement between two segmentations, that also considers the agreement occurring by chance. The metric ranges from 0 to 1, with 0 meaning a complete disagreement between both segmentations and 1 means a complete overlapping.

$$\kappa = \frac{P_A - P_E}{1 - P_E}$$

$$P_A = Accuracy = \frac{TP + TN}{P + N}$$

$$P_E = \left(\frac{P}{P + N} \cdot \frac{\hat{P}}{\hat{P} + \hat{N}}\right) + \left(\frac{N}{P + N} \cdot \frac{\hat{N}}{\hat{P} + \hat{N}}\right)$$

Chapter 5

Results

The results obtained by the unsupervised segmentation algorithms evaluated in this M.Sc. Thesis, in combination with the proposed preprocessing and postprocessing pipelines are shown in Table 5.1. The results are obtained through the BRATS 2013 evaluation web page provided for the Segmentation Challenge of MICCAI 2013 conference. The Table 5.1 shows the average results for the 10 patients of the BRATS 2013 test set, grouped by the unsupervised algorithms, the tumour subcompartments and the figures of merit used to assess the quality of the segmentation.

Classifier	Dice			PPV			Sensitivity			Kappa
	comp	core	enh	comp	core	enh	comp	core	enh	
K-means	0.69	0.49	0.57	0.66	0.48	0.68	0.76	0.57	0.51	0.98
Fuzzy K-means	0.70	0.46	0.39	0.73	0.47	0.51	0.71	0.54	0.35	0.98
GMM	0.69	0.60	0.55	0.63	0.60	0.64	0.78	0.68	0.55	0.98
GHMRF	0.72	0.62	0.59	0.68	0.58	0.67	0.81	0.75	0.60	0.98

Table 5.1: Summary of results of the unsupervised segmentation algorithms evaluated in the study, in combination with the proposed preprocessing and postprocessing pipelines. The results are the average of the 10 patients of the BRATS 2013 test set, grouped by the different unsupervised algorithms, tumour subcompartments and figures of merit. *comp* refers to complete tumour subcompartment, *core* refers to tumour core subcomparment and *enh* refers to enhancing tumour subcomparment.

In most cases GHMRF rises as the best algorithm in combination with the proposed preprocessing and postprocessing pipelines. Dice coefficient, as well as sensitivity, reveals that GHMRF retrieves the best results in all the subcompartment segmentations. PPV is the only statistic in which other algorithms achieve better results, indicating that GHMRF may be slightly inferior in precision than the other methods. Also, it could be seen that, regardless of the non-supervised algorithm used, the complete tumour subcompartment always achieves the highest scores with respect to the other subcompartments, primarily due to the presence of the edema tissue in this subcompartment, which often may be the bigger and easiest class to identify.

Table 5.2 shows the published ranking of the BRATS 2013 competition and the position of the proposed unsupervised segmentation method in its best configuration, i.e GHMRF in combination with the preprocessing and postprocessing. As it can be seen, our method

reaches the 7th position in the competition against mostly supervised approaches. Ranks in the table are defined by sorting each column in descending order and computing the average of the positions for each user in each column.

Pos	User	Dice			PPV			Sensitiviy			Kappa
		comp	core	enh	comp	core	enh	comp	core	enh	
1	N. Tustison	0.87	0.78	0.74	0.85	0.74	0.69	0.89	0.88	0.83	0.99
2	R. Meier	0.82	0.73	0.69	0.76	0.78	0.71	0.92	0.72	0.73	0.99
3	S. Reza	0.83	0.72	0.72	0.82	0.81	0.70	0.86	0.69	0.76	0.99
4	L. Zhao	0.84	0.70	0.65	0.80	0.67	0.65	0.89	0.79	0.70	0.99
5	N. Cordier	0.84	0.68	0.65	0.88	0.63	0.68	0.81	0.82	0.66	0.99
6	J. Festa	0.72	0.66	0.67	0.77	0.77	0.70	0.72	0.60	0.70	0.98
7	**This work**	**0.72**	**0.62**	**0.59**	**0.68**	**0.58**	**0.67**	**0.81**	**0.75**	**0.60**	**0.98**
8	S. Doyle	0.71	0.46	0.52	0.66	0.38	0.58	0.87	0.70	0.55	0.98

Table 5.2: Ranking of the *BRATS* 2013 Segmentation Challenge with the position occupied by our proposed unsupervised segmentation framework (with the GHMRF algorithm). *comp* refers to complete tumour subcompartment, *core* refers to tumour core subcomparment and *enh* refers to enhancing tumour subcomparment.

.

Finally, several examples of segmentations achieved by the different unsupervised segmentation algorithms obtained through the proposed method are shown in Figure 5.1.

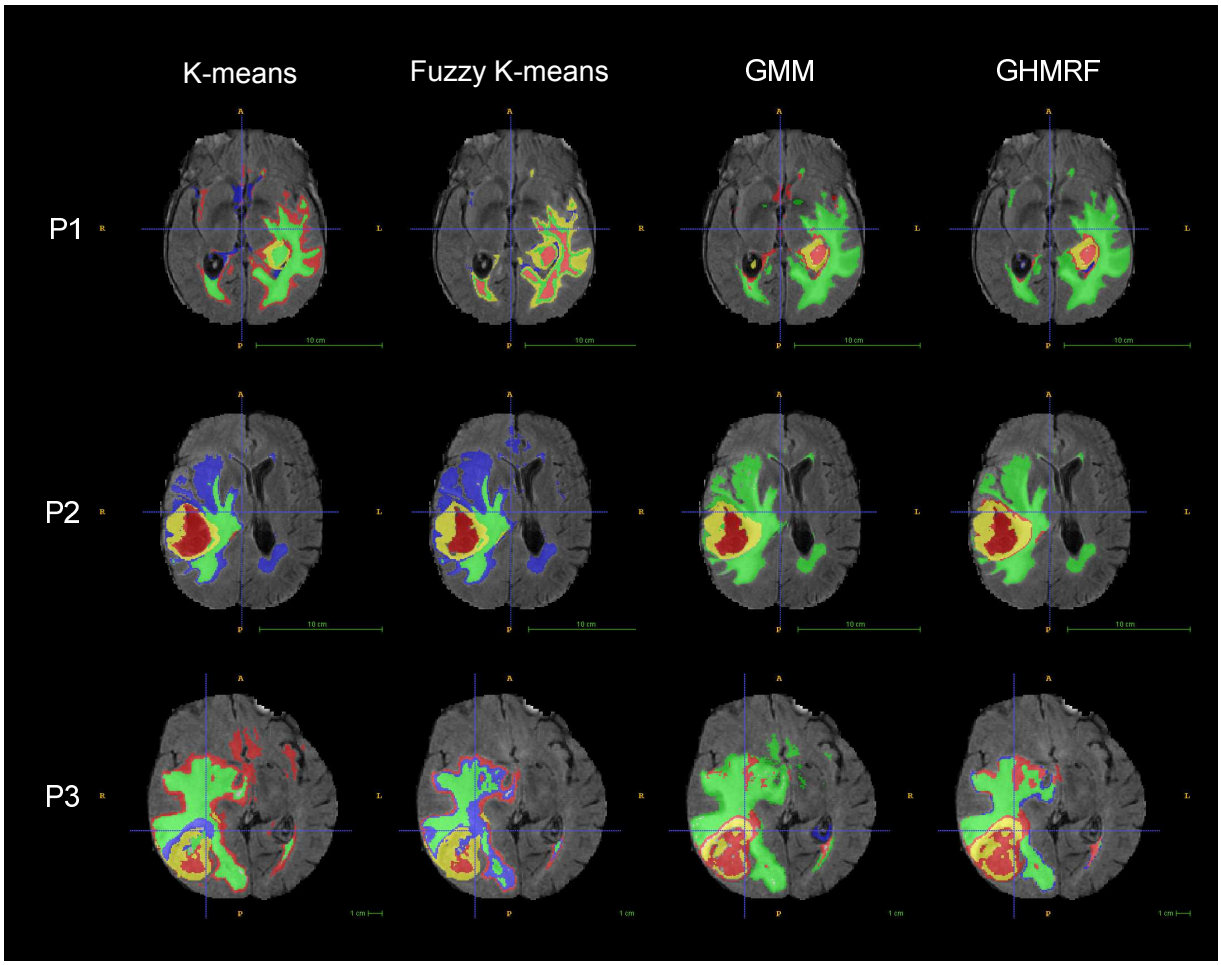

Figure 5.1: Examples of final segmentations (preprocess and postprocess included) of 3 patients of BRATS dataset computed by different non-supervised algorithms.

Chapter 6

Discussion

The proposed unsupervised brain tumour segmentation method is confirmed as a viable alternative for GBM segmentation, as it has demonstrated to achieve competitive and comparable results in a public reference brain tumour dataset such as the BRATS 2013 Challenge (See Table 5.2). The method is able to obtain competitive results without any prior knowledge or manual expert labelling, thus overcoming the limitations of the supervised approaches such as the time-consuming and biased task of retrieving a training dataset. Furthermore, the method provides a general mechanism to automatically isolate the tumoral tissues in the brain, to address the problems associated with the biological interpretability of the unsupervised results. This mechanism can be also extrapolated to other pathologies, as it adopts an *a contrario* approach, by identifying the normal tissues and then isolating the abnormal classes that represent the pathology.

The proposed unsupervised segmentation method comprises four stages: MRI preprocessing, Feature extraction and dimensionality reduction, Unsupervised voxel classification and Automatic tumour classes isolation. Concerning the preprocessing stage, consolidated state of the art techniques that provide efficient solutions to enhance the information of the MR images are employed. However, some preprocessing techniques are primarily oriented to non-pathological brains. This is the case of bias field correction. In our experiments, we found that the estimation of the magnetic field inhomogeneities with the N4 algorithm presented problems primarily with FLAIR sequences. The hyper-intensity shown in the FLAIR sequence by the edema was confused frequently with inhomogeneities of the magnetic field, thereby reducing its intensity and sometimes removing it from the image. In order to overcome this problem we reduced the number of iterations of the algorithm to 10 iterations at each scale, to remove as much inhomogeneities as possible, while keeping the intensities of the lesion. Such solution assumes a non optimal removal of the magnetic field inhomogeneities, but allows to save the information contained in the lesion area, which becomes more important to the brain tumour segmentation. However, it is mandatory to develop MRI preprocessing techniques suited to pathological brains, to enhance the images as much as possible, while avoiding the removal of useful information to classify/segment the pathology.

Several unsupervised classification algorithms are evaluated to assess its pros and cons. Attending to the restrictiveness and constraints imposed to the probabilistic models that underlies the unsupervised algorithms, an evident hierarchy becomes apparent. These constraints implies both positive and negative effects, such as the efficiency in the algorithm inference or the quality and fidelity of the model fitted to the data. K-means is the

most restrictiveness algorithm in terms of the class-conditional probabilistic model. From a Gaussian generative standpoint, K-means assumes an equal and identity covariance matrix for all classes, equal prior probabilities and hard assignment of each observation to a class. Also, an iid assumption between the observations to be classified is considered, which together derives in a more simplistic, not necessarily worse, classifier. In contrast, GHMRF is the less restrictive algorithm in terms of the class-conditional probabilistic model (also GMM clustering), but it imposes a prior structure to the data that constraints the inference process of the algorithm. Such prior structure introduces statistical dependencies between adjacent variables of the probabilistic model, that penalizes neighbouring voxels with different classes. Hence, this structured prior aims to model the self similarity presented in the images, leading the algorithm to a more homogeneous segmentation than the non-structured classification techniques. Finally, GMM clustering is the non-structured version of GHMRF, while Fuzzy K-means is a particular case of GMM, which assumes an equal and identity covariance matrix for all classes.

Therefore, it is expected that the less restrictive algorithms in terms of class-conditional probability model are likely to achieve better results, based on the hypothesis that such algorithms learn a model for the classes that better fit the data to be classified (a more realistic model). Moreover, algorithms that introduces mechanisms to model the self similarity of the images are also expected to retrieve better results based on the hypothesis that they exploits the information redundancy of the images. Table 5.1 confirms such hypothesis. The results shown in Table 5.1 are the metrics provided by the BRATS 2013 evaluation web page grouped by the unsupervised algorithms and the tumour subcompartments. GHMRF rises as the best algorithm in almost all the metrics. Only the PPV reveals that other algorithms may achieve a slightly better precision. Figure 5.1 also corroborates the hypothesis. GHMRF segmentation leads to more homogeneous segmentations, which is consistent with the pathological standpoint and hence with the manual labelling that an expert radiologist will provide.

Note that differences between the GHMRF and the K-means segmentations of the peritumoral and distal areas of the tumour can be observed (for example in the P2 patient of Figure 5.1). The K-means segmentation shows a division of this area into two classes, which are finally labelled as edema (green color class) and non-enhancing tumour (blue color class). Based on the definition of non-enhancing tumour, it seems clear that the area classified as non-enhancing tumour in K-means and Fuzzy K-means does not correspond with this tissue, and probably refers to edema. The reason by which this area is partially misclassified into non-enhancing tumour comes from the automatic tumour classes isolation stage. As proposed in Section 4.4 (Subsection 4.4.3) we enforce the system to return 4 or less classes, depending on their statistical distribution similarities. Hence, we assume that similar classes are merged in the 4.4.3 step, returning a set of final classes different enough to be considered a different tissue related to the BRATS labelling. In this sense, in the K-means and Fuzzy K-means segmentations, we were required to set a distinct label for each of the 4 classes returned by the method, thereby partially misclassifying the distal area as non-enhancing tumour. In contrast, although GHMRF also initially divided this area into two classes, the method found both regions similar enough to be merged, hence determining the whole area as an homogeneous class, which is more coherent than the K-means or Fuzzy K-means segmentations.

However, it is worth noting that this is an example of the ability of the unsupervised

approaches to explore or provide natural clusters of data that can be useful to make new clinical hypothesis. Following the P2 segmentation example, it can be seen that the hyper-intensity area of the FLAIR sequence classified as edema by the GHMRF is not homogeneous. Such area was initially separated by all clustering techniques into two classes, therefore revealing evident differences inside the edema, that may have interesting clinical interpretations. Such capabilities provide an added value to the segmentation systems that arise from an unsupervised learning standpoint.

Finally, a key point that should be discussed refers to the viability of supervised segmentation approaches to real clinical routine. The difficulties in the acquisition of manual labelled ground truths and the recalibration of the models when the data sources experiment changes (for example different hospitals or new MR protocols) severely affects the performance of supervised approaches. Furthermore, supervised algorithms are trained to classify the already well-known clinical knowledge, which allows to reduce the manual cost of the tumour segmentation, but does not provide new knowledge that helps physicians to make alternative clinical hypothesis. Conversely, unsupervised approaches inherits the exploratory ability of clustering techniques, which are able to provide physicians a guide to interpret sub-segmentations based on natural groupings of data.

The analysis of functional Magnetic Resonance Imaging (fMRI) such as PWI is an example where clustering techniques may play an important role. fMRI is a technique for measuring the brain activity by detecting changes in the blood oxygenation and flow in response to neural activity. The ability of unsupervised algorithms to explore segmentations where different tissues determined in the anatomical MRI analysis share a similar behaviour in the fMRI approach should be explored. Such similarities may provide an insight of the evolution of the different tissues and hence provide useful information to make early decisions that improve the treatment of the disease.

Functional MR imaging is rising as the future of brain tumour diagnosis due to its ability to reveal biomarkers related to the behaviour of the tissues instead of their anatomical properties. These biomarkers might predict relevant information such as the tumour growing direction and its evolution. In this sense, we believe that research efforts should be aligned with MR functional imaging requirements by providing powerful systems that covers its clinical purposes. We plan to extend our unsupervised segmentation method to the analysis and segmentation of PWI.

Chapter 7

Concluding remarks and future work

7.1 Conclusion

The present M.Sc. Thesis provides a method based on the ML discipline to solve the brain tumour segmentation problem. The first contribution of this M.Sc. Thesis concerns the design and implementation of a robust methodology for unsupervised brain tumour segmentation. We provide a fully automated method able to achieve accurate results comparable to supervised approaches, but avoiding the tedious, time-consuming and biased task of manual expert labelling. The second major contribution refers to a statistical postprocessing method able to robustly identify which classes in a brain tumour segmentation corresponds to normal tissues. Hence, the method allows to automatically isolate the pathological classes in the brain that belong to abnormal tissues. Finally, we present a comprehensive evaluation of several unsupervised segmentation algorithms attending to its structured and non-structured condition. We use a public real brain tumour dataset in order to make a comparison between the state-of-the-art techniques in supervised segmentation and the unsupervised method proposed in this M.Sc. Thesis.

The conclusions extracted from this M.Sc. Thesis are:

- The results of the comprehensive evaluation through the public BRATS 2013 brain tumour dataset show that the proposed unsupervised segmentation method provides accurate and coherent segmentations, similar than the manual labelling provided by an expert radiologist. Hence, it is confirmed as a viable method for brain tumour segmentation.

- It is mandatory to develop MRI preprocessing techniques suited to pathological brains, to enhance these images as much as possible, while avoiding the removal of useful information to classify or segment the disease.

- Structured classification, specifically MRF, provides a statistical language to define probabilistic models that represent dependencies between random variables, that better suit to image segmentation problems, as it can model the self similarity of the images.

- The proposed postprocessing approach to improve the biological interpretability of the unsupervised results is able to identify and isolate the pathological classes of a segmentation that correspond to abnormal tissues in the brain. Hence, the

method provides clinicians an unsupervised segmentation of the whole brain, with the possibility of automatically identify the abnormal classes of the segmentation.

- A well designed unsupervised segmentation method can yield comparable results to supervised approaches, without the need of prior manual expert labelling. Thus, it becomes a viable alternative to supervised approaches for real clinical application.

7.2 Future work

Some of the future lines of research directly related to the results of this M.Sc. Thesis are:

- The future of GBM treatment points to the analysis of functional imaging such as Perfusion Weighted Images (PWI). The biological information provided by these MR acquisitions may indicate the behaviour of the tumour, such as neoangiogenesis, and its evolution. It is mandatory to include such information in the segmentation process to try to identify the growing areas of the tumour and its direction.

- Monitoring the evolution of a GBM tumour leads to the acquisition of several MR images during different temporal moments. Considering this temporal information may improve the segmentation of the tumour, and even may provide a prediction of the tumour growing direction in a posterior temporal moment.

- The feature extraction and dimensionality reduction techniques proposed in this M.Sc. Thesis may be improved through the computation of the Second order texture features (Haralick texture features [49]) and non-linear dimensionality reduction techniques such as non-linear PCA. The implementation of these texture features and the evaluation of the improvement through the non-linear reduction methods are currently carrying out.

- The robust characterization of brain tumour tissues through biological signatures based on the clinical information recovered for the patient is a challenging and ambitious goal that should be tackled. Such standardized biological signatures may provide a common reference framework for studying tumoral tissues.

Glossary

Acronyms

AI	Artificial Intelligence
ANN	Artificial Neural Networks
ANTS	Advanced Normalization Tools
BRATS	BRAin Tumour Segmentation
CNS	Central Nervous System
CDSS	Clinical Decision Support System
CSF	Cerebro-Spinal Fluid
CRF	Conditional Random Fields
DTI	Diffusion Tensor Imaging
E	Expectation
EM	Expectation-Maximization
fMRI	functional Magnetic Resonance Imaging
FLAIR	FLuid Attenuated Inversion Recovery
GBM	Glioblastoma Multiforme
GM	Gray Matter
GMM	Gaussian Mixture Model
GHMRF	Gaussian Hidden Markov Random Field
HAC	Hierarchical Agglomerative Clustering
HG	High Grade
HMRF	Hidden Markov Random Fields
ICBM	International Consortium for Brain Mapping
i.e	id est (that is)
iid	independent and identically distributed

ICM	Iterated Conditional Modes
LG	Low Grade
M	Maximization
MICCAI	Medical Image Computing and Computer-Assisted Intervention
ML	Machine Learning
MR	Magnetic Resonance
MRF	Markov Random Field
MRI	Magnetic Resonance Imaging
NLM	Non Local Means
NMR	Nuclear Magnetic Resonance
PCA	Principal Component Analysis
PD	Proton Density
PR	Pattern Recognition
PPV	Positive predictive value
PR	Pattern Recognition
PWI	Perfusion Weighted Images
RF	Radio Frequency
ROBEX	RObust Brain EXtraction
ROI	Region Of Interest
SOM	Self Organizing Map
SVM	Support Vector Machines
TE	Echo Time
TR	Repetition Time
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
UPV	Universitat Politècnica de València
WM	White Matter

Bibliography

[1] T. A. Dolecek, J. M. Propp, N. E. Stroup, and C. Kruchko, "Cbtrus statistical report: Primary brain and central nervous system tumors diagnosed in the united states in 2005–2009," *Neuro-Oncology*, vol. 14, no. 5, pp. 1–49, 2012.

[2] A. von Deimling, *Gliomas*, vol. 171 of *Recent Results in Cancer Research*. Springer Berlin Heidelberg, 2009.

[3] P. Y. Wen, D. R. Macdonald, D. A. Reardon, T. F. Cloughesy, A. G. Sorensen, E. Galanis, J. DeGroot, W. Wick, M. R. Gilbert, A. B. Lassman, C. Tsien, T. Mikkelsen, E. T. Wong, M. C. Chamberlain, R. Stupp, K. R. Lamborn, M. A. Vogelbaum, M. J. van den Bent, and S. M. Chang, "Updated response assessment criteria for high-grade gliomas: Response assessment in neuro-oncology working group," *Journal of Clinical Oncology*, vol. 28, no. 11, pp. 1963–1972, 2010.

[4] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of mri-based medical image analysis for brain tumor studies," *Physics in Medicine and Biology*, vol. 58, no. 13, p. R97, 2013.

[5] K. L. Wagstaff, *Intelligent Clustering with instance-level constraints*. PhD thesis, Cornell University, 2002.

[6] R. Stupp, W. P. Mason, M. J. van den Bent, M. Weller, B. Fisher, M. J. Taphoorn, K. Belanger, A. A. Brandes, C. Marosi, U. Bogdahn, J. Curschmann, R. C. Janzer, S. K. Ludwin, T. Gorlia, A. Allgeier, D. Lacombe, J. G. Cairncross, E. Eisenhauer, and R. O. Mirimanoff, "Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma," *New England Journal of Medicine*, vol. 352, no. 10, pp. 987–996, 2005.

[7] A. Mang, J. Schnabel, W. Crum, M. Modat, O. Camara-Rey, C. Palm, G. Caseiras, H. Jäger, S. Ourselin, T. Buzug, and D. Hawkes, "Consistency of parametric registration in serial mri studies of brain tumor progression," *International Journal of Computer Assisted Radiology and Surgery*, vol. 3, no. 3-4, pp. 201–211, 2008.

[8] E. D. Angelini, O. Clatz, E. Mandonnet, E. Konukoglu, L. Capelle, and H. Duffau, "Glioma dynamics and computational models: A review of segmentation, registration, and in silico growth algorithms and their clinical applications," *Current Medical Imaging Reviews*, vol. 3, no. 4, pp. 262–276, 2007.

[9] S. Wang and R. M. Summers, "Machine learning and radiology," *Medical Image Analysis*, vol. 16, no. 5, pp. 933 – 951, 2012.

[10] N. Gordillo, E. Montseny, and P. Sobrevilla, "State of the art survey on {MRI} brain tumor segmentation," *Magnetic Resonance Imaging*, vol. 31, no. 8, pp. 1426 – 1438, 2013.

[11] H. Cai, R. Verma, Y. Ou, S. koo Lee, E. Melhem, and C. Davatzikos, "Probabilistic segmentation of brain tumors based on multi-modality magnetic resonance images," in *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pp. 600–603, April 2007.

[12] R. Verma, E. I. Zacharaki, Y. Ou, H. Cai, S. Chawla, S.-K. Lee, E. R. Melhem, R. Wolf, and C. Davatzikos, "Multiparametric tissue characterization of brain neoplasms and their recurrence using pattern classification of {MR} images," *Academic Radiology*, vol. 15, no. 8, pp. 966 – 977, 2008.

[13] S. Ruan, S. Lebonvallet, A. Merabet, and J. Constans, "Tumor segmentation from a multispectral mri images by using support vector machine classification," in *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pp. 1236–1239, April 2007.

[14] S. Ruan, N. Zhang, Q. Liao, and Y. Zhu, "Image fusion for following-up brain tumor evolution," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pp. 281–284, March 2011.

[15] M. Tayel and M. Abdou, "A neuro-difference fuzzy technique for automatic segmentation of region of interest in medical imaging," in *Radio Science Conference, 2006. NRSC 2006. Proceedings of the Twenty Third National*, vol. 0, pp. 1–7, March 2006.

[16] T. R. Jensen and K. M. Schmainda, "Computer-aided detection of brain tumor invasion using multiparametric mri," *Journal of Magnetic Resonance Imaging*, vol. 30, no. 3, pp. 481–489, 2009.

[17] C.-H. Lee, S. Wang, A. Murtha, M. Brown, and R. Greiner, "Segmenting brain tumors using pseudo–conditional random fields," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008* (D. Metaxas, L. Axel, G. Fichtinger, and G. Székely, eds.), vol. 5241 of *Lecture Notes in Computer Science*, pp. 359–366, Springer Berlin Heidelberg, 2008.

[18] S. Bauer, L.-P. Nolte, and M. Reyes, "Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011* (G. Fichtinger, A. Martel, and T. Peters, eds.), vol. 6893 of *Lecture Notes in Computer Science*, pp. 354–361, Springer Berlin Heidelberg, 2011.

[19] K. Popuri, D. Cobzas, A. Murtha, and M. Jägersand, "3d variational brain tumor segmentation using dirichlet priors on a clustered feature set," *International Journal of Computer Assisted Radiology and Surgery*, vol. 7, no. 4, pp. 493–506, 2012.

[20] L. R. Schad, S. Blüml, and Z. I, "A survey of mri-based medical image analysis for brain tumor studies," *Physics in Medicine and Biology*, vol. 58, no. 13, p. R97, 2013.

[21] L. M. Fletcher-Heath, L. O. Hall, D. B. Goldgof, and F. Murtagh, "Automatic segmentation of non-enhancing brain tumors in magnetic resonance images," *Artificial Intelligence in Medicine*, vol. 21, no. 1–3, pp. 43 – 63, 2001. Fuzzy Theory in Medicine.

[22] J. Nie, Z. Xue, T. Liu, G. S. Young, K. Setayesh, L. Guo, and S. T. Wong, "Automated brain tumor segmentation using spatial accuracy-weighted hidden markov random field," *Computerized Medical Imaging and Graphics*, vol. 33, no. 6, pp. 431 – 441, 2009.

[23] Y. Zhu, G. S. Young, Z. Xue, R. Y. Huang, H. You, K. Setayesh, H. Hatabu, F. Cao, and S. T. Wong, "Semi-automatic segmentation software for quantitative clinical brain glioblastoma evaluation," *Academic Radiology*, vol. 19, no. 8, pp. 977 – 985, 2012.

[24] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm," *Medical Imaging, IEEE Transactions on*, vol. 20, pp. 45–57, Jan 2001.

[25] C. Vijayakumar, G. Damayanti, R. Pant, and C. Sreedhar, "Segmentation and grading of brain tumors on apparent diffusion coefficient images using self-organizing maps," *Computerized Medical Imaging and Graphics*, vol. 31, no. 7, pp. 473 – 484, 2007.

[26] R. Damadian, "Tumor detection by nuclear magnetic resonance," *Science*, vol. 171, no. 3976, pp. 1151–1153, 1971.

[27] P. C. Lauterbur, "Image formation by induced local interactions: Examples employing nuclear magnetic resonance," *Nature*, vol. 242, no. 5394, pp. 190–191, 1973.

[28] P. C. Lauterbur, "Magnetic resonance zeugmatography," *Pure and Applied Chemistry*, vol. 40, no. 1-2, pp. 149–157, 1974.

[29] A. Drevelegas and N. Papanikolaou, *Imaging Modalities in Brain Tumors*. Imaging of Brain Tumors with Histological Correlations, Springer Berlin Heidelberg, 2011.

[30] W. Pope and C. Hessel, "Response assessment in neuro-oncology criteria: Implementation challenges in multicenter neuro-oncology trials," *American Journal of Neuroradiology*, vol. 32, no. 5, pp. 794–797, 2011.

[31] H. Gudbjartsson and S. Patz, "The rician distribution of noisy mri data," *Magnetic Resonance in Medicine*, vol. 34, no. 6, pp. 910–914, 1995.

[32] I. Diaz, P. Boulanger, R. Greiner, and A. Murtha, "A critical review of the effects of de-noising algorithms on mri brain tumor segmentation," in *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, pp. 3934–3937, 2011.

[33] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model & Simulation*, vol. 4, pp. 490–530, 2005.

[34] J. V. Manjón, P. Coupé, L. Martí-Bonmatí, D. L. Collins, and M. Robles, "Adaptive non-local means denoising of mr images with spatially varying noise levels," *Journal of Magnetic Resonance Imaging*, vol. 31, no. 1, pp. 192–203, 2010.

[35] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain {MRI} registration," *NeuroImage*, vol. 46, no. 3, pp. 786 – 802, 2009.

[36] B. Avants, C. Epstein, M. Grossman, and J. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Medical Image Analysis*, vol. 12, no. 1, pp. 26 – 41, 2008. Special Issue on The Third International Workshop on Biomedical Image Registration – {WBIR} 2006.

[37] C. Fennema-Notestine, I. B. Ozyurt, C. P. Clark, S. Morris, A. Bischoff-Grethe, M. W. Bondi, T. L. Jernigan, B. Fischl, F. Segonne, D. W. Shattuck, R. M. Leahy, D. E. Rex, A. W. Toga, K. H. Zou, and G. G. Brown, "Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: Effects of diagnosis, bias correction, and slice location," *Human Brain Mapping*, vol. 27, no. 2, pp. 99–113, 2006.

[38] B. Dogdas, D. W. Shattuck, and R. M. Leahy, "Segmentation of skull and scalp in 3-d human mri using mathematical morphology," *Human Brain Mapping*, vol. 26, no. 4, pp. 273–285, 2005.

[39] J. Iglesias, C.-Y. Liu, P. Thompson, and Z. Tu, "Robust brain extraction across datasets and comparison with publicly available methods," *Medical Imaging, IEEE Transactions on*, vol. 30, pp. 1617–1634, Sept 2011.

[40] J. Sled, A. Zijdenbos, and A. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in mri data," *Medical Imaging, IEEE Transactions on*, vol. 17, no. 1, pp. 87–97, 1998.

[41] N. Tustison, B. Avants, P. Cook, Y. Zheng, A. Egan, P. Yushkevich, and J. Gee, "N4itk: Improved n3 bias correction," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 6, pp. 1310–1320, 2010.

[42] E. Plenge, D. Poot, W. Niessen, and E. Meijering, "Super-resolution reconstruction using cross-scale self-similarity in multi-slice mri," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013* (K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, eds.), vol. 8151 of *Lecture Notes in Computer Science*, pp. 123–130, Springer Berlin Heidelberg, 2013.

[43] J. V. Manjón, P. Coupé, A. Buades, D. L. Collins, and M. Robles, "Mri superresolution using self similarity and image priors," *International Journal of Biomedical Imaging*, vol. 2010, p. 11, 2010.

[44] F. Rousseau, "A non-local approach for image super-resolution using intermodality priors," *Medical Image Analysis*, vol. 14, no. 4, pp. 594–605, 2010.

[45] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *Image Processing, IEEE Transactions on*, vol. 18, no. 1, pp. 36–51, 2009.

[46] J. V. Manjón, P. Coupé, A. Buades, V. Fonov, D. L. Collins, and M. Robles, "Non-local {MRI} upsampling," *Medical Image Analysis*, vol. 14, no. 6, pp. 784–792, 2010.

[47] A. Kassner and R. Thornhill, "Texture analysis: A review of neurologic mr imaging applications," *American Journal of Neuroradiology*, vol. 31, no. 5, pp. 809–816, 2010.

[48] S. Ahmed, K. M. Iftekharuddin, and A. Vossough, "Efficacy of texture, shape, and intensity feature fusion for posterior-fossa tumor segmentation in mri," *Trans. Info. Tech. Biomed.*, vol. 15, pp. 206–213, Mar. 2011.

[49] R. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, vol. 67, pp. 786–804, May 1979.

[50] L. Van Gool, P. Dewaele, and A. Oosterlinck, "Texture analysis anno 1983," *Computer Vision, Graphics, and Image Processing*, vol. 29, pp. 336–357, 1985.

[51] T. Reed and J. Dubuf, "A review of recent texture segmentation and feature extraction techniques," *CVGIP: Image Understanding*, vol. 57, no. 3, pp. 359 – 372, 1993.

[52] M. Chantler, *The effect of variation in illuminant direction on texture classification.* PhD thesis, Dept. Computing and Electrical Engineering, Heriot-Watt University, 1994.

[53] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, (Philadelphia, PA, USA), pp. 1027–1035, Society for Industrial and Applied Mathematics, 2007.

[54] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[55] C. Bishop, *Neural Networks for Pattern Recognition.* Clarendon Press, 1995.

[56] S. Lloyd, "Least squares quantization in pcm," *Information Theory, IEEE Transactions on*, vol. 28, pp. 129–137, Mar 1982.

[57] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, (Berkeley, Calif.), pp. 281–297, University of California Press, 1967.

[58] R. Duda, P. Hart, and D. Stork, *Pattern classification.* Pattern Classification and Scene Analysis: Pattern Classification, Wiley, 2001.

[59] J. C. Dunn, "A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters," *Journal of Cybernetics*, vol. 3, no. 3, pp. 32–57, 1973.

[60] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms.* Norwell, MA, USA: Kluwer Academic Publishers, 1981.

[61] J. M. Hammersley and P. Clifford, "Markov fields on finite graphs and lattices." 1971.

[62] N. Komodakis and G. Tziritas, "Approximate labeling via graph cuts based on linear programming," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, pp. 1436–1453, Aug 2007.

[63] N. Komodakis, G. Tziritas, and N. Paragios, "Performance vs computational efficiency for optimizing single and dynamic mrfs: Setting the state of the art with primal-dual strategies," *Comput. Vis. Image Underst.*, vol. 112, pp. 14–29, Oct. 2008.

[64] C. Saez, M. Robles, and J. Garcia-Gomez, "Stability metrics for multi-source biomedical data based on simplicial projections from probability distribution distances," *Statistical Methods in Medical Research*, vol. Forthcoming, 2014.

List of Figures

List of Tables