UNIVERSIDAD AUTONOMA DE MADRID

Escuela Politécnica Superior

Departamento de Ingeniería Informática

Doctorado en Ingeniería Informática y Telecomunicación

THESIS

Intelligent Analysis of Cerebral Magnetic Resonance Images: Extracting Relevant Information from Small Datasets

Author

Ania Benítez Sánchez del Campo

Advisors

Manuel Sánchez-Montañés Isla

Sebastián Cerdán García-Esteller

June 2017

Acknowledgements

A la Agencia Española de Cooperación Internacional para el Desarrollo por otorgarme la beca de investigación que me permitió dar inicio a este trabajo.

A mis tutores Dr. Sebastián Cerdán García-Esteller y Dr. Manuel Sánchez-Montañés, Isla por la certera dirección, el constante interés en el desarrollo exitoso de este trabajo y el financiamiento a través de sus proyectos. Al Dr. Cerdán, especialmente, por acogerme en su grupo y darme la oportunidad de crecer profesionalmente; por su maestría y entrega a la ciencia, su confianza e invaluable ayuda. Al Dr. Sánchez-Montañés por implicarse tanto en el desarrollo de este trabajo y aportar sus excelentes ideas e inagotable empeño en intentar mejorar más y más los resultados.

A la Dra. Arisbel Cerpa Naranjo por recomendarme y ponerme en contacto con el Dr. Cerdán y estar siempre al pendiente del progreso de esta tesis.

A la Dra. Pilar López Larrubia, quien siempre ha tenido respuestas a mis preguntas, por sus atinadas sugerencias y ser para mí un ejemplo de mujer de ciencia y persona.

Al profesor Dr. Luis Lago Fernández por ofrecerme sus opiniones y colaboración.

A mis compañeros del Instituto de Investigaciones Biomédicas "Alberto Sols" (IIB) CSIC/UAM: Patricia Sánchez García, María José Guillen Gómez, Viviana Negri, Teresa Navarro Hernanz, María Rodríguez Martínez, Irene Guadilla Gómez y a todos con los que he coincidido durante estos años de trabajo. Es un equipo genial, del que me he sentido afortunada de formar parte, por su profesionalidad, cercanía y entusiasmo, siempre dispuestos a ayudar. En especial agradezco a Blanca Lizarbe Serra y Daniel Calle Hernández, mis compañeros de laboratorio junto a los que comencé la investigación doctoral, por inspirarme con su responsabilidad y compromiso en el trabajo, por su compañerismo y buenos consejos. A Blanca Lizarbe Serra, Rocío Pérez Carro, Ana Amor e Irene Guadilla Gómez por compartir conmigo sus datos experimentales. A Javier Pérez García, "el dibujante", por su trabajo esmerado e infinita paciencia.

A mi familia por compartir todos mis desvelos y mis éxitos, porque aun estando lejos físicamente, siempre logran hacerse sentir muy cerca. En especial a mis padres, por inculcarme la pasión por el estudio y apoyarme siempre, incondicionalmente, en todas las facetas de mi vida. También a mi abuela Martha, por su inmenso cariño y la dedicación y complicidad en todos mis propósitos.

A Pablo Morales Mombiela por tanto y tan buen amor; por compartir mis objetivos y escuchar mis preocupaciones. Agradezco mucho a Pablo sus soluciones,

revisiones, redacciones y críticas constructivas, sin conformarse nunca con un 'porque sí' y animarme a buscar siempre mejores argumentos. Gracias por todo y por el tiempo dedicado.

A las de "los zapaticos de rosa", por estar siempre y darme, con su cariño, alas de mariposa.

A los amigos de "aquí" y de "allá", por animarme a seguir hasta al final.

A todos los que me han ayudado de alguna manera.

¡Muchísimas gracias!

We can only see a short distance ahead, but we can see plenty there that needs to be done.

Alan Turing

Abstract

Machine learning methods applied to medical imaging are becoming important tools for the analysis and diagnose of patients. The huge availability of repositories from multimodal imaging has favored the development of systems that learn to extract relevant features and construct predictive models from enormous amounts of data, for example, deep learning methods. However, the analysis of imaging datasets from smaller number of subjects, as normally gathered in preclinical and clinical research environments, has received considerably less attention. In this thesis, we implement a variety of advanced computational tools to overcome this problem, supporting robust analysis of Magnetic Resonance Imaging (MRI) from applications that involve small numbers of subjects. We illustrate these approaches analyzing automatically datasets obtained from functional MR images of the cerebral regulation of appetite in rodents and humans, and from functional and structural MR images from glioma development in animal models and humans. The proposed methods evolved from the idea of considering each voxel from the image dataset as a pattern, rather than from the conventional notion of considering each image as a pattern.

Chapter 1 describes the motivations supporting these developments, including the objectives proposed, the general structure of the document and the contributions of this research. **Chapter 2** provides an updated introduction to the state of the art in MRI, the conventional image pre-processing methods, and the most useful machine learning algorithms and their MRI applications. **Chapter 3** presents the experimental design, and image pre-processing steps as applied to the datasets from appetite regulation and glioma development. **Chapter 4** implements new supervised learning algorithms for the analysis of MRI datasets as obtained from a small number of subjects. We illustrate this approach, presenting first the Fisher Maps methodology, which allow for the non-invasive and comprehensive visualization of the integral cerebral appetite circuitry, through the automatic analysis of Diffusion Weighted Image (DWI) datasets. This methodology is expanded to the classification of complete images, using the combined predictions obtained from all their pixels. **Chapter 5** proposes new unsupervised learning algorithm for the analysis of MRI datasets and illustrates its performance with synthetic data and datasets of brain tumoral studies and glioma development. Finally, **Chapter 6** summarizes the main conclusions, providing ample avenues for the continuation of this work.

In summary, the present thesis provides new useful approach for the automatic extraction of relevant information from the analysis of MRI datasets in contexts where small datasets are available, using advanced supervised and unsupervised artificial intelligence algorithms. The proposed methods may be easily extended to other paradigms or pathologies, and even to, alternative imaging modalities.

Resumen

Los métodos de *machine learning* aplicados a imágenes médicas se están convirtiendo en potentes herramientas para el análisis y diagnóstico de pacientes. La alta disponibilidad de repositorios de imágenes de diferentes modalidades ha favorecido el desarrollo de sistemas que aprenden a extraer características relevantes y construyen modelos predictivos a partir de grandes cantidades de información, por ejemplo, los métodos de *deep learning*. Sin embargo, el análisis de conjuntos de imágenes provenientes de un menor número de sujetos, como es el caso de las imágenes adquiridas en entornos de investigación clínica y pre-clínica, ha recibido considerablemente menos atención. El objetivo de esta tesis es implementar un conjunto de herramientas avanzadas para resolver este problema, permitiendo el análisis robusto de Imágenes de Resonancia Magnética (MRI por sus siglas en inglés) cuando se dispone de pocos sujetos de estudio. En este contexto, las herramientas propuestas se emplean para analizar de manera automática conjuntos de datos obtenidos de imágenes funcionales de MR del cerebro en estudios de regulación del apetito en roedores y humanos, y de imágenes funcionales y estructurales de MR de desarrollos tumorales en modelos animales y humanos. Los métodos propuestos se derivan de la idea de considerar cada voxel del conjunto de imágenes como un patrón, en lugar de la noción convencional de considerar cada imagen como un patrón.

El **Capítulo 1** describe la motivación de esta tesis, incluyendo los objetivos propuestos, la estructura general del documento y las contribuciones de esta investigación. El **Capítulo 2** contiene una introducción actualizada del estado del arte en MRI, los procedimientos más usados en el pre-procesamiento de imágenes, y los algoritmos de *machine learning* más útiles y sus aplicaciones en MRI. El **Capítulo 3** presenta el diseño experimental y los pasos de pre-procesamiento aplicados a los conjuntos de datos de regulación de apetito y desarrollo tumoral. El **Capítulo 4** implementa nuevos métodos de aprendizaje supervisados para el análisis de conjuntos de datos de MRI obtenidos de un conjunto pequeño de sujetos. Se ilustra este enfoque presentando primero la metodología *Fisher Maps*, que permite la visualización cuantitativa y no invasiva de la circuitería cerebral del apetito, mediante el análisis automático de Imágenes Ponderadas en Difusión (DWI por sus siglas en inglés). Esta metodología se extiende a la clasificación de imágenes completas combinando las predicciones obtenidas de cada píxel. El **Capítulo 5** propone un nuevo algoritmo de aprendizaje no supervisado, ilustrando su desempeño sobre datos sintéticos y datos provenientes de estudios de tumores cerebrales y crecimiento tumoral. Por último, en el **Capítulo 6** se resumen las principales conclusiones de este trabajo y se plantean amplias vías para su desarrollo futuro.

En resumen, esta tesis presenta un nuevo enfoque capaz de trabajar en contextos con baja disponibilidad de sujetos de estudio, proponiendo algoritmos de aprendizaje

supervisado y no supervisado. Estos métodos pueden ser fácilmente generalizados a otros paradigmas o patologías, e incluso, a distintas modalidades de imágenes.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter outlines the motivations underlying the implementation of this PhD. thesis, the objectives proposed, the general structure of the PhD. document, and the current list of publications derived from it.

1.1 Motivation

Discovered in 1948 simultaneously by the groups of Felix Bloch at Stanford, and Edward Purcell at Harvard [Bloch, 1946, Purcell et al., 1946], the magnetic resonance phenomenon has currently reached impressive applications in Physics and Chemistry. The possibility to obtain magnetic resonance images (MRI) of live subjects, proposed by Lauterbur and Mansfield [Lauterbur, 1973, Mansfield and Maudsley, 1977], and Damadian [Damadian et al., 1974], expanded further its frontiers to biological sciences and opened a new era in biomedicine in general, and in diagnostic medicine, in particular.

An essential part of this progress has been achieved through the development of optimized computational tools to process the large amounts of data generated by magnetic resonance experiments, extracting only the relevant physiological or diagnostic information. Computational developments supporting MRI applications are in continuous evolution, as new MRI applications develop, the frontiers of diagnostic medicine expand and new disease signatures are discovered. However, a frequent limitation to these developments is that the educated opinion of a specialist is required to advise on the most plausible interpretation of MR images, precluding complete exploitation of the MRI advantages in automatic, operator free, environments. Recent progress in artificial intelligence methods may improve significantly this situation, facilitating robust computational tools to improve our knowledge on the mechanisms underlying disease, identifying therapeutic responses automatically, earlier than detected by conventional specialists, or providing support to diagnostic decisions using objective, quantitative, criteria.

Recently, the huge availability in the web of data repositories [NITRC, 2007, FNLCR, 2014] from multimodal imaging, has favored the implementation of deep learning methods [Shen et al., 2017], to be able to interpret the large amounts of data gathered. However, the analysis of imaging data obtained from significantly

smaller number of subjects, as normally generated in preclinical and clinical research environments for specific projects, has received considerably less attention. On these grounds, I proposed in this thesis the implementation a variety of advanced computational tools to support recent applications of MRI in environments involving a small number of subjects.

1.2 Objectives

The general objective of this thesis is to implement novel, automatic, supervised and unsupervised image processing methods to favor analysis and interpretation of biomedical image datasets from a small number of subjects. We followed this general purpose along two main lines; a) Development of classification systems to analyze efficiently these imaging datasets; and b) Implementation of automatic systems to recognize new patterns contained in MR imaging datasets, disclosing previously unknown features.

More specific objectives are:

1. To provide an overview of the used magnetic resonance methods and review critically the different methods and techniques used in pre-processing and automatic analysis of images in the literature, framing the context and state of the art for this work.

2. To develop advanced supervised learning methods to classify MR imaging datasets from small number of subjects, investigated with different MR imaging sequences. We propose to create classifiers considering individual pixels, assigning them a degree of pertinence to a predefined class, and classify the complete image using the ensemble of predictions for all its pixels.

3. To apply these algorithms to real MR imaging datasets from rodents and humans, including; cerebral appetite stimulation or glioma grading and dimorphism.

4. To implement a novel clustering algorithm to analyze jointly various datasets, identifying both, invariant and specific patterns.

5. To apply this novel clustering algorithm to synthetic data and MRI datasets.

1.3 Structure of the thesis

This thesis contains six chapters:

Chapter 1: Introduction. This chapter presents the motivation, objectives, structure, and current contributions of this PhD thesis.

Chapter 2: State of the Art. The second chapter introduces the context for this thesis. I address briefly the fundamentals of MR imaging, including a brief summary of previous applications to obesity and cancer research. Furthermore, I elaborate in more detail on the essentials of machine learning methods that have been applied on MR image analysis, including supervised and unsupervised learning algorithms.

Chapter 3: Experimental datasets and pre-processing. In the third chapter, I describe the different experimental image datasets used later in the computer analysis including mainly; the appetite datasets from rodents and humans, and those derived from animal models of glioma and human patients diagnosed with glioma. The image pre-processing steps implemented for each dataset are commented.

Chapter 4: Supervised learning. I provide a new supervised methodology to automatically explore MR images in fourth chapter, implementing two classification methods applied either to the appetite impulse in the brain of mice, rats and humans or to rodent glioma models and human patients diagnosed with glioma.

Chapter 5: Unsupervised learning. In the fifth chapter, I propose a new clustering algorithm based on the Expectation-Maximization (EM) method, describing its mathematical background and the results obtained by applying it to synthetic data and to MRI datasets.

Chapter 6: Conclusions and future perspectives. Finally, I present the main conclusions that can be derived from this research and propose some interesting lines of future work.

1.4 Contributions

Some of the results of this PhD. thesis have been already published in international journals or presented at international conferences. These publications, as well as two related software packages are briefly described in this section.

International journals

1. Benítez, A., Lizarbe, B., López-Larrubia, P., Lago-Fernández, L., Cerdán, S., and Sánchez-Montañés, M.(2017). Fisher maps for the analysis of DWI datasets from cerebral appetite activation. *NMR in Biomedicine*. (Submitted

in May 2017).

Impact factor: 2.983 (JCR, 2015) [Q1, Subject area: Radiology, Nuclear Medicine and Imaging].

Summary: We propose a new protocol for the analysis of Diffusion Weighted Imaging (DWI) datasets of cerebral activation by appetite in mice, rats and humans inspired in Fisher´s Linear Discriminant Analysis (LDA), and compare the results with those provided by the classical mono-exponential diffusion model. DWI datasets from the same individual were acquired from a brain section containing the hypothalamus and coplanar cortico-limbic structures, under fed and fasted conditions. Mono-exponential fittings revealed significant Apparent Diffusion Coefficient (ADC) decreases through the brain upon fasting, but rigorous parameter estimations imposed the rejection of considerable numbers of pixels. The Fisher Maps approach introduced here avoided pixel rejections and provided a representation of the DWI dataset as a pixel map of the "Appetite Index" (AI), a parameter revealing the appetite score of every pixel. AI maps visualized, comprehensively, the stimulation of the cerebral appetite circuitry, independent of the limitations of previous knowledge considerations, diffusion models or statistical parametric mapping methods. Both AI and ADC methodologies yielded consistent results, demonstrating simultaneous activation of hypothalamic and cortico-limbic structures, albeit with different relative intensities, in rodents and humans. In particular, rodents and humans showed similar hypothalamic activations, but rodents depicted more intense activations of cortico-limbic structures than humans, for similar fasting periods. In summary, Fisher Maps complement adequately the classical ADC maps approach using the same datasets and may become easily extended to other activation paradigms or alternative neuroimaging methods.

Contribution: Experimental design, algorithm design and implementation, pre-processing and analysis of imaging data, and part of the writing of the manuscript. This contribution is related to Chapter 3 (Section 3.1) and Chapter 4 (Sections 4.2 and 4.4.1).

2. Lizarbe, B., Benítez, A., Sánchez-Montañés, M., Lago-Fernández, L., Garcia-Martin, M.L., López-Larrubia, P., and Cerdán, S. (2013). Imaging hypothalamic activity using diffusion weighted magnetic resonance imaging in the mouse and human brain. *NeuroImage*, 64:448–457.

Impact factor: 6.132 (JCR, 2013) [Q1].

Google Scholar Citations: 10 (Jun 14^{th}, 2017).

Summary: Hypothalamic appetite regulation is a vital homeostatic process underlying global energy balance in animals and humans, its disturbances resulting in feeding disorders with high morbidity and mortality. The objective evaluation of appetite remains difficult, very often restricted to indirect measurements of food intake and body weight. We report here, the direct,

non-invasive visualization of hypothalamic activation by fasting using diffusion weighted magnetic resonance imaging, in the mouse brain as well as in a preliminary study in the human brain. The brain of fed or fastedmice or humanswere imaged at 7 or 1.5 Tesla, respectively, by diffusionweighted magnetic resonance imaging using a complete range of b values ($10 < b < 2000 s/mm^2$). The diffusion weighted image data sets were registered and analyzed pixel by pixel using a biexponential model of diffusion, or a model-free Linear Discriminant Analysis approach. Biexponential fittings revealed statistically significant increases in the slow diffusion parameters of the model, consistent with a neurocellular swelling response in the fasted hypothalamus. Increased resolution approaches allowed the detection of increases in the diffusion parameters within the Arcuate Nucleus, Ventromedial Nucleus and Dorsomedial Nucleus. Independently, Linear Discriminant Analysis was able to classify successfully the diffusion data sets from mice and humans between fed and fasted states. Present results are consistent with increased glutamatergic neurotransmission during orexigenic firing, a process resulting in increased ionic accumulation and concomitant osmotic neurocellular swelling. This swelling response is spatially extendable through surrounding astrocytic networks until it becomes MRI detectable. Present findings open new avenues for the direct, non-invasive, evaluation of appetite disorders and other hypothalamic pathologies helping potentially in the development of the corresponding therapies.

Contribution: Experimental design of model-independent approach, algorithm design and implementation of model-independent approach, pre-processing and analysis of imaging data using model-independent approach, and part of the writing of the manuscript. Both first authors contributed equally. This contribution is related to Chapter 3 (Section 3.1) and Chapter 4 (Sections 4.3 and 4.4.2).

3. Lizarbe, B., Benitez, A., Peláez-Brioso, G. A., Sánchez-Montañés, M., López-Larrubia, P., Ballesteros, P., and Cerdán, S. (2013). Hypothalamic metabolic compartmentation during appetite regulation as revealed by magnetic resonance imaging and spectroscopy methods. *Frontiers in Neuroenergetics*(*), Review Article, Volume 5, Article 6, pp. 1-14.

   Downloads: 1598, Citations: 8 (Jun 14^{th}, 2017) [Q2 in SJR, 2013].

   Google Scholar Citations: 13 (Jun 14th, 2017).

   (*) This specialty journal has closed, now its topics have been included in the new publication Neuroenergetics, Nutrition and Brain Health section of Frontiers in Neuroscience (Impact factor: 3.566, JCR 2016).

   Addicionally, this work was also published in a book: Cerdan, S. (Ed.). (2015). Transcellular Cycles underlying Neurotransmission. Frontiers Media SA. doi: 10.3389/978-2-88919-654-8. The book was consulted 77,671 times and this paper was consulted 11,590 times (Jun 14^{th}, 2017).

Summary: We review the role of neuroglial compartmentation and transcellular neurotransmitter cycling during hypothalamic appetite regulation as detected by Magnetic Resonance Imaging (MRI) and Spectroscopy (MRS) methods. We address first the neurochemical basis of neuroendocrine regulation in the hypothalamus and the orexigenic and anorexigenic feed-back loops that control appetite. Then we examine the main MRI and MRS strategies that have been used to investigate appetite regulation. Manganese-enhanced magnetic resonance imaging (MEMRI), Blood oxygenation level-dependent contrast (BOLD), and Diffusion-weighted magnetic resonance imaging (DWI) have revealed Mn^{2+} accumulations, augmented oxygen consumptions, and astrocytic swelling in the hypothalamus under fasting conditions, respectively. High field ^{1}H magnetic resonance in vivo, showed increased hypothalamic myo-inositol concentrations as compared to other cerebral structures. ^{1}H and ^{13}C high resolution magic angle spinning (HRMAS) revealed increased neuroglial oxidative and glycolytic metabolism, as well as increased hypothalamic glutamatergic and GABAergic neurotransmissions under orexigenic stimulation. We propose here an integrative interpretation of all these findings suggesting that the neuroendocrine regulation of appetite is supported by important ionic and metabolic transcellular fluxes which begin at the tripartite orexigenic clefts and become extended spatially in the hypothalamus through astrocytic networks becoming eventually MRI and MRS detectable.

Contribution: Critical reading of literature and part of the writing of the manuscript. This contribution is related to Chapter 2.

4. Lizarbe, B., Benitez, A., L. Lago-Fernández, L., M. Sánchez-Montañés, M., López-Larrubia, P., and Cerdán, S. (2011). Intelligent image analysis of diffusion weighted data sets: a new tool for functional imaging. In *International Workshop on Combinatorial Image Analysis*, 6636:9-12. Part of the Lecture Notes in Computer Science (LNCS) book series, Springer, Berlin, Heidelberg.

Downloads: 488 (Jun 14^{th}, 2017).

Summary: Obesity is a pandemic syndrome underlying the most prevalent causes of death and disability in developed countries including atherosclerosis, ischemic episodes and cancer. Obesity results from an imbalance in global energy metabolism, ultimately caused by disturbances in the neuroendocrine control of appetite in the hypothalamus and leading to an uncompensated feeding/fasting balance [Schwartz et al., 2000]. On these grounds, the non invasive detection of hypothalamic activation by food under healthy or diseased conditions entails considerable interest for the diagnosis and treatment of obesity and other food intake disorders as anorexia or bulimia.

Contribution: Critical literature review, algorithm design and implementation, writing first draft. This contribution is related to Chapter 4.

International conferences

1. Benítez, A., Lizarbe, B., Guadilla, I., López-Larrubia, P., Lago-Fernández, L.F., Sánchez-Montañés, M., and Cerdán, S. (2016). Asymmetrical cerebral response to appetite in mice. In Proceedings of *24th Annual Meeting & Exhibition of International Society for Magnetic Resonance in Medicine*. Singapore, May 7-13.

   Summary: We investigate the global cerebral response to appetite in mice by functional Diffusion Weighted Imaging, implementing two independent, but complementary, methods of data analysis; a) Model-free classification algorithm and b) Biexponential diffusion parameter fittings. The model-free approach allowed the pixel by pixel calculation of appetite index maps, used to classify the brain between fed and fasted conditions. Biexponential model fittings allowed the calculation of diffusion parameter maps, revealing significant increases in diffusion parameters through the whole fasted brain. Both methods detected an asymmetric cerebral response to appetite with the right cerebral hemisphere becoming more responsive.

   Contribution: Algorithm design and implementation, pre-processing and analysis of imaging data, and abstract writing. This contribution is related to Chapter 4.

2. Benítez, A., Peláez-Brioso, G., Borges, A., López-Larrubia, P., Cerdán, S., and Sánchez-Montañés, M. (2014). An alternative approach for the automatic prediction of therapy response from MRI data sets in small cohorts of experimental high grade gliomas. In Proceedings of *Joint Annual Meeting ISMRM-ESMRMB*. Milano, Italy, May 10-16.

   Summary: Clinical management of High Grade Gliomas (HGG) often involves antiangiogenic therapies using monoclonal antibodies, mainly against VGEF. Responses to this treatment are highly heterogeneous, with similar cohorts of patients behaving as "responders" or "non-responders". It would be then very useful to discriminate by non-invasive criteria the two populations as early as possible, to tailor treatment accordingly. Magnetic Resonance Imaging (MRI) is presently one of the most important non-invasive methods to investigate and diagnose HGG. Similarly, the automatic classification of medical images into different pathological categories is currently an active area of research in developing diagnostic support environments. A common problem to both areas is the relatively small size of experimental observations available to establish robust classifications. Although the sample size is small, the data set about each subject is normally very big: several MRIs types, slices, pixels and many times, longitudinal days of treatment. Thus, automatic processing of all this information adequately entails vital relevance. Here we propose an alternative protocol to classic approaches, implementing the automatic classification by selecting patterns and attributes from MRI data set using Pixel-Based

Classification (P-BC). Moreover, we present a pilot study on the discriminant power of the method as compared to the classical approach in predicting the outcome of the treatment to anti-VGEF therapy in mice bearing implanted GL621 tumors.

Contribution: Algorithm design and implementation, pre-processing and analysis of imaging data, and abstract writing. This contribution is related to Chapter 3 (Section 3.2) and Chapter 4 (Section 4.3).

3. Benítez, A., Arias-Ramos, N., Pacheco-Torres, J., Cerdan, S., and López-Larrubia, P. (2013). A quantitative analysis of glioma response to hypoxia targeted therapy using intelligent image processing in a rat model. In Proceedings of *30th Annual Scientific Meeting of European Society for Magnetic Resonance in Medicine and Biology (ESMRMB)*. Toulouse, Francia, October 3-5.

   Summary: Oxygen tension (pO_2) performs a very important role in the appropriate working of tissues and organs as well as in the correct developing of an organism. It has been seen that hypoxia (low oxygen concentration) has a fundamental role in the evolution of different pathologies situations like tumors, strokes, isquemias, etc. In addition, is has been shown that hypoxia also have a negative influence in resistance to therapies and poor patient outcome. Several treatments have been specifically developed to modulate hypoxia in order to improve therapy success, as breathing a gas with high oxygen content during radiotherapy among others. Responses to this treatments are highly heterogeneous, because tumors are not homogeneous among patients and at intratumoral level, so with similar cohorts of patients behaving as "responders" or "non-responders". It would be then very useful to discriminate by non invasive criteria the two populations as early as possible, to tailor treatment accordingly. Here we report a multivariate analysis approach on the discriminant power of MRI to predict the irradiation therapy success in hyperoxic conditions by using a high grade gliomas (HGG) rat model.

   Contribution: Algorithm design and implementation, pre-processing and analysis of imaging data, and abstract writing. This contribution is related to Chapter 4 (Section 4.3).

4. Benítez, A., Peláez-Brioso, G., Borges, A., López-Larrubia, P., Lago-Fernández, L., Cerdán, S., and Sánchez-Montañés, M. (2013). My Therapy Response: A novel software tool to predict the response to antiangiogenic therapy in mouse glioma models. In Proceedings of *30th Annual Scientific Meeting of European Society for Magnetic Resonance in Medicine and Biology (ESMRMB)*. Toulouse, Francia, October 3-5.

   Summary: Clinical management of high grade gliomas (HGG) often involves antiangiogenic therapies using monoclonal antibodies, mainly against VGEF. Responses to this treatment are highly heterogeneous, with similar cohorts of patients behaving as "responders" or "non-responders". It would be then very

useful to discriminate by non invasive criteria the two populations as early as possible, to tailor treatment accordingly. Here we provide a software package to predict in vivo and automatically the treatment outcome of mice bearing implanted Gl621 tumors to anti-VGEF therapy, using DWI and T_{2w} data sets.

Contribution: Algorithm design, software implementation, and abstract writing. This contribution is related to Chapter 3 (Section 3.2) and Chapter 4 (Section 4.3).

5. Benítez, A., Borges, A., López-Larrubia, P., Lago-Fernández, L., Sánchez-Montañés, M., and Cerdán, S. (2012). A quantitative analysis of glioma response to antiangiogenic therapy using intelligent image processing. In Proceedings of *29th Annual Scientific Meeting of European Society for Magnetic Resonance in Medicine and Biology (ESMRMB)*. Lisboa, Portugal, October 4-6.

   Summary: Clinical management of high grade gliomas (HGG) often involves antiangiogenic therapies using monoclonal antibodies, mainly against VGEF. Responses to this treatment are highly heterogenous, with similar cohorts of patients behaving as "responders" or "non-responders". It would be then very useful to discriminate by non invasive criteria the two populations as early as possible, to tailor treatment accordingly. Here we report on a multivariate analysis approach on the discriminant power of T_{2w} and DWI, or their combination to predict the treatment outcome of mice bearing implanted Gl621 tumors to anti-VGEF therapy.

   Contribution: Algorithm design and implementation, pre-processing and analysis of imaging data, and abstract writing. This contribution is related to Chapter 3 (Section 3.2) and Chapter 4 (Section 4.3).

6. Benítez, A., Peláez-Brioso, G., Lizarbe, B., Lago-Fernández, L., López-Larrubia, P., Cerdán, S., and Sánchez-Montañés, M. (2012). My appetite: A novel software tool to identify appetite disorders. In Proceedings of *29th Annual Scientific Meeting of European Society for Magnetic Resonance in Medicine and Biology (ESMRMB)*. Lisboa, Portugal, October 4-6.

   Summary: Obesity is a pandemic syndrome, one of the most prevalent and morbid diseases in developed countries. It is caused by a disorder of the feeding behavior, which is controlled by intrahypothalamic orexigenic and anorexigenic neurons. The diagnosis of obesity is often made by body measurements including weight, but it would be more precise to assess its development directly at the hypothalamic level. We provide here a software package to evaluate in vivo and automatically the "fed" or "fasted" state of the hypothalamus in the mouse and human brains, using Diffusion Weighted Imaging data sets with multiple b values. The program may be useful to assess hypothalamic function in obesity and other eating disorders.

   Contribution: Algorithm design, software implementation, and abstract writing. This contribution is related to Chapter 3 (Section 3.1) and Chapter 4

(Section 4.3).

7. Benítez, A., Lizarbe, B., Lago-Fernández, L., López-Larrubia, P., Cerdán, S. and Sánchez-Montañés, M. (2011). Automatic classification of fed and fasted mouse brain using multivariate analysis of diffusion weighted images. In Proceedings of *28th Annual Scientific Meeting of European Society for Magnetic Resonance in Medicine and Biology (ESMRMB)*. Leipzig, Alemania, October 6-8.

   Summary: Statistical Parametric Mapping (SPM) [Friston et al., 2007] is one the most popular techniques allow for the automatic analysis of functional neuroimaging, but it requires extensive pre-processing of images and does not directly solve the problem of classifying images into pathological categories. In this study, we propose a new technique that directly and automatically classifies diffusion weighted images (DWI) of mouse brain into fed and fasted states, without the need of extensive pre-processing stage.

   Contribution: Algorithm design and implementation, pre-processing and analysis of imaging data, and abstract writing. This contribution is related to Chapter 3 (Section 3.1) and Chapter 4 (Section 4.3).

8. Benítez, A., Lizarbe, B., Lago-Fernández, L., López-Larrubia, P., Cerdán, S. and Sánchez-Montañés, M. (2011). An automatic protocol to detect the fed and fasted brain using multivariate analysis of diffusion weighted data sets. In Proceedings of *19th Annual Meeting & Exhibition of International Society for Magnetic Resonance in Medicine*. Montreal, Canada, May 7-13.

   Summary: The automatic classification of medical images into different pathological categories is currently an active area of research. In the context of functional neuroimaging, a variety of techniques allow for the automatic analysis of the acquired images, the most popular being Statistical Parametric Mapping (SPM) [Friston et al., 2007]. It is aimed at the generation of parametric maps which are able to reveal specific cerebral activation events as obtained by functional imaging methods. However, SPM requires extensive pre-processing of images in order to perform the statistical analysis, in many cases involving the intervention of an expert. Moreover, it does not directly solve the problem of classifying images into pathological categories, but it just identifies local areas probably related to the case of interest or pathology. In this work, we propose a new technique that is able to resolve, directly, the classification of images into two different classes without the need of a pre-processing stage, and illustrate the performance of the method with the classification of diffusion weighted MR images from the brain of fed and fasted mice. Using a collection of diffusion weighted images (DWI), it became possible to ascribe the images to a fed or fasted animal in 100% of the cases.

   Contribution: Algorithm design and implementation, pre-processing, and analysis of imaging data, and abstract writing. This contribution is related to Chapter 3 (Section 3.1) and Chapter 4 (Section 4.3).

The abstract acceptance rate for ISMRM/ESMRMB congresses varies year-to-year, typically between 70-80% (http://www.ismrm.org/about/ampc-faq, Jun 15^th, 2017).

Software packages

Additionally, two software packages were implemented for the Department of Experimental Models of Human Disease in the "Alberto Sols" Biomedical Research Institute CSIC/UAM.

1. *My Therapy Response*: Software tool to predict the response to antiangiogenic therapy in mouse glioma models.

   My Therapy Response provides a protocol for the prediction of therapy response to antiangiogenic therapy in mouse glioma models, based on non-invasive image parameters as T_{2w} and DWI. This software package is implemented in MATLAB (The Mathworks, Natick, US) environment and is able to accept DWI and T_{2w} datasets from mouse brains, providing automatically a prediction for the new subject in a scale of "response rate" 0 (non-responders) - 100 (responders). Features include: Data input, Data Exploration, Train new model, Classify a subject, and Export report.

   Specific Hardware Requirements, Onsite: Matlab Version 7.4.0.287 (R2007a) or superior.

2. *My Appetite*: Software tool to identify appetite disorders.

   My Appetite is implemented in MATLAB (The Mathworks, Natick, US) environment and is able to accept DWI datasets from mouse or human brains, providing automatically a classification of the new image in a scale of "appetite index" 0 (fed) - 100 (fasted). The user may enter his own external dataset into the application, use preeexisting dataset already configured in the software package or a combination of both. Features include: Data input, Data Exploration, Train new model, Classify a subject, and Export report.

   Specific Hardware Requirements, Onsite: Matlab Version 7.4.0.287 (R2007a) or superior.

Chapter 2

State of the art

This chapter provides an overview of the state of the art in the different topics covered by the present PhD thesis. I begin with a brief description on the fundamentals of Magnetic Resonance Imaging (MRI) including the basis of nuclear magnetism, imaging principles and sequences, image contrast and most recent applications to cerebral appetite disorders and cancer. Then, I review the most common image preprocessing strategies including slice selection, brain extraction, bias field correction, intensity normalization and registration. Finally, I address the most useful image analysis protocols, including both, conventional and machine learning approaches. In the latter, I emphasize in the description of supervised and unsupervised learning algorithms and their applications to the analysis of MR images.

2.1 Magnetic resonance imaging and applications

2.1.1 Nuclear magnetism

Magnetic Resonance is based in the interaction between the magnetic moments μ of nuclei and an external magnetic field B [Abragam, 1961, Ernst et al., 1987].

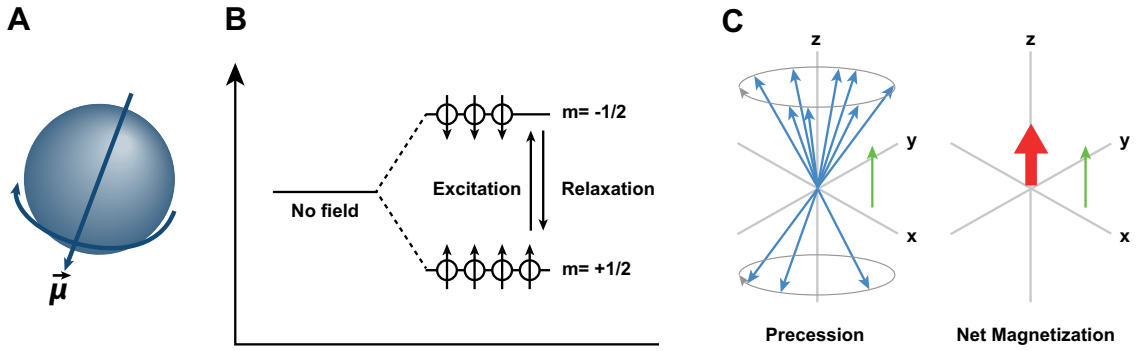

Figure 2.1: Magnetic moment of a nucleus and magnetization properties. A: Magnetic moment of a rotating proton (left). B: Energy levels of nuclear orientation and corresponding spin populations. Nuclear excitation and relaxation. C: Orientation and precession of an ensemble of magnetic moments in the presence of an external magnetic field B_0 (left). Spins tend to orient in the less energetic, B_0 parallel direction, resulting in a net magnetization $\vec{M}$ (right).

In the case of water, which represents approximately 80% of the weight in mammals, individual hydrogen atoms have a nuclear spin quantum number of m=1/2, indicating that when an external magnetic field B_0 is applied, two possible energetically different states are created (Figure 2.1 B). These are associated to magnetic moments $\mu = \pm 1/2$, and energies $E = \pm\frac{\gamma h}{2\pi} m B_0$, where γ and h represent the gyromagnetic ratio ($rad \cdot s^{-1} T^{-1}$) and the Plank constant ($J \cdot s$), respectively. The lower energy state $\mu = +1/2$ (E_{low}) has the magnetic moment parallel to the B_0 field, while the higher energy state $\mu = -1/2$ (E_{high}) has a magnetic moment antiparallel to the B_0 field. It is possible to induce transitions between adjacent nuclear energy sates by applying a short radiofrequency (RF) pulse B_1, perpendicularly to the B_0 direction;

$$\Delta E = E_{\text{high}} - E_{\text{low}} = \gamma h B_0 \tag{2.1}$$

The excitation process promotes spins from the "low" to the "high" energy level, disturbing transiently the thermal equilibrium distribution of spins. Thermal equilibrium of spins tends to recover after the pulse, by releasing the energy accumulated to the surroundings (or *lattice*), in the relaxation process known as *spin-lattice relaxation*. In this process, the equilibrium longitudinal magnetization is recovered with a time constant termed T_1 (longitudinal relaxation time). T_1 depends on the rotational and translational dynamics of the surrounding molecules of the lattice and thus, the T_1 of water will vary among different tissues, depending on their characteristic magnetic properties [Gore and Kennan, 1999, Hendrick and Raff, 1992]. Simultaneously, the transverse magnetization decays because its magnetic moments dephase as a result of their mutual interaction. This process is known as T_2 relaxation, which differs also among tissues [Gore and Kennan, 1999, Hendrick and Raff, 1992]. During T_2 decay, no energy is transferred from the nuclei to the tissue, and nuclei in the "high" and "low" states exchange energy with each other. The main contribution to T_2 decay in biological tissues comes from the magnetic fields of the neighboring protons from the same water molecule, or from different protons of larger macromolecules, which reorient more slowly, promoting shorter T_2 relaxation. Notably, the experimentally observed decrease in transverse magnetization occurs more rapidly than expected for pure T_2 relaxation alone, due to the presence of inevitable magnetic field inhomogeneity's that accelerate the dephasing between protons. This shorter transversal relaxation time is termed $T_2{}^*$.

The magnetic resonance phenomenon can also be explained with a vector interpretation [Ernst et al., 1987] (Figure 2.1 C). The slight excess in spin population of the low energy level, results in the appearance of a macroscopic magnetic moment in the sample, oriented in the direction of the B_0 field. During excitation with an RF pulse, the magnetic moment experiences a torque, becoming rotated to the xy plane. When the RF excitation ceases, the magnetic moment recovers exponentially to the original equilibrium position, increasing magnetization in the z direction (Longitudinal relaxation T_1) or dissipating the magnetization in the xy plane (Transversal magnetization T_2) [Gore and Kennan, 1999].

When incorporating the relaxation effects to the time dependence of magnetization described in equation 2.1, it results

$$\frac{d\vec{M}}{dt} = \gamma \vec{M} x \vec{B} - \frac{M_x \vec{1} + M_y \vec{j}}{T_2} - \frac{(M_z + M_0)\vec{k}}{T_1} \tag{2.2}$$

that describes the evolution of magnetization with time, and is known as the Bloch equation [Bloch, 1946].

Most MR techniques manipulate tissue magnetization repeatedly by using a train of RF excitation pulses separated by a *repetition time* (TR). Under steady state conditions, the longitudinal magnetization recovers approximately to a fraction $1 - e^{\frac{TR}{T_1}}$ of its equilibrium value between RF pulses. Pulse sequences manipulating longitudinal magnetization recovery are known as progressive saturation or inversion recovery sequences [Bydder et al., 1998]. Different pulse sequences may be implemented to enhance contrast in T_2. In this case, a 180°RF refocusing pulse is applied at a time τ after the 90°pulse, originating an echo after 2τ, known as the echo time (TE). Under these conditions, the amplitude of the transverse magnetization depends on TE value and T_2, and is proportional to $e^{\frac{-TE}{T_2}}$. Sequences using this mechanism are known as *spin echo* (SE) sequences [Bradley et al., 1999].

2.1.2 Imaging nuclear magnetization

Figure 2.2 illustrates the basic principles of Magnetic Resonance Imaging [Hashemi et al., 2012, McRobbie et al., 2003]. Briefly, obtaining an MRI image from an object requires localizing the magnetization of every pixel of the object in a 3D space. This is done in modern MRI instruments (Figure 2.2 A), by codifying the information in the x, y and z directions using magnetic field gradients. In the presence of a magnetic field gradient, the resonance frequency v_i of a hydrogen nucleus as described in 2.1, is modified to

$$v_i = \gamma / h(B_0 + G_i) \tag{2.3}$$

where G_i is the intensity of the magnetic field gradient in the i_{th} direction. This makes different the resonance frequency of every nucleus, depending on its relative position within the gradient, and allows to identity its position in the x, y or z coordinates by its resonance frequency. Similarly, the presence of a magnetic field gradient alters the phase of the precession Φ_i, in a manner that nuclei in different positions of the gradient will precess with different phases

$$\Phi_i = \gamma / h(\phi + G_i) \tag{2.4}$$

Therefore, it is possible to codify the position of a nucleus the x, y, z coordinates by having different frequencies in two directions and a different phase in the third one, as generated by three orthogonal field gradients. These magnetic field gradients

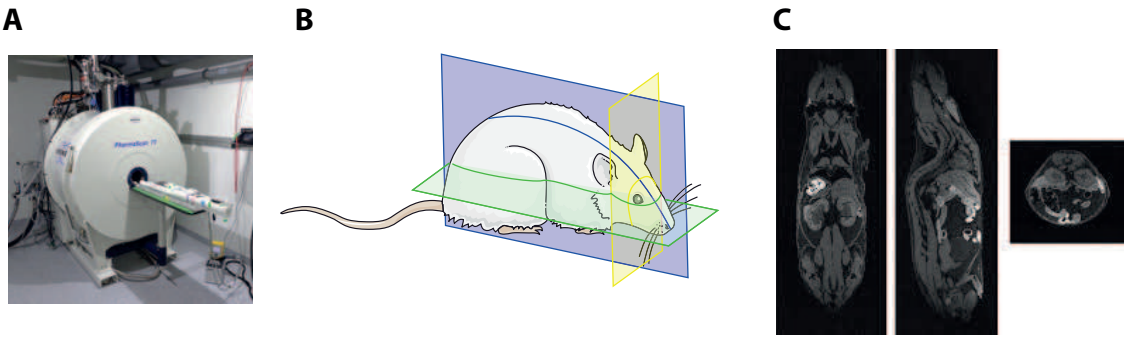

Figure 2.2: Fundamentals of MRI. A: 9,4 Tesla magnet from a representative pre-clinical MRI scanner. B: Orthogonal magnetic fields created by appropriate coils encode the frequency and phase of every pixel in a manner that can be retrieved by appropriate pulse sequences. C: Representative coronal (left), sagittal (center) and axial (right) sections through a representative mouse body.

are provided by three orthogonal coils surrounding the object (Figure 2.2 B), which allow to obtain a virtual section across the object in any direction of space depending on the relative strengths of the three gradients [Matwiyoff and Brooks, 1999] (Figure 2.2 C).

Briefly, the imaging process begins by activating pulsed, orthogonal, magnetic field gradients and RF pulses appropriately [Hashemi et al., 2012]. The first applied gradient, the slice-selection gradient (G_{s1}, Figure 2.3 A), is switched on while applying the RF pulse, in order to restrict the MR signal to a selected slice across the object. Shortly after the RF pulse has been turned off, a phase-encoding gradient G_p is pulsed perpendicular to the slice-selection gradient, thus making the frequencies from the selected slice to depend further on spatial position (now on the phase encoding direction). The final step of spatial encoding is the application of a frequency-encoding gradient (the "read out" gradient G_r) along the third orthogonal direction, at the same time that the signal is received. This modifies the Larmor frequencies again and creates proton columns with identical frequencies, but different phases, along the "read out" direction. In summary, the combination of G_s, G_p and G_r gradients enables to codify the position of every pixel in the z, y and x coordinates, respectively.

In general, magnetic resonance images are obtained from voltage measurements acquired in the time domain (Figure 2.3 B), but represented as intensities in the frequency domain [McRobbie et al., 2003, Wood and Wehrli, 1999] (Figure 2.3 C). The mathematical tool that expresses the relationship between these two different ways of presenting the same information is the Fourier Transformation. Fourier transformations decompose the voltage signals from every voxel or pixel, into a distribution of cosine and sine waves of different frequencies. More specifically, the

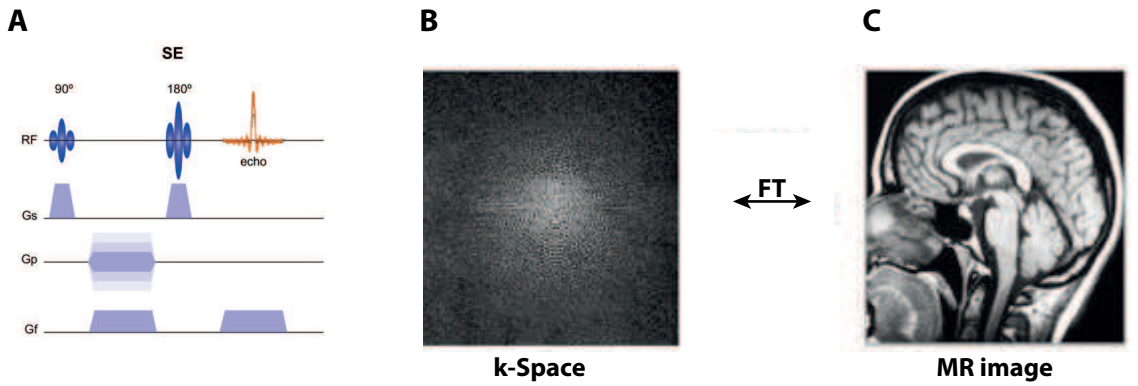

Figure 2.3: The spin-echo sequence. A: Schematic representation of a spin-echo (SE) imaging experiment. During the activation of the 90º RF pulse, a slice-selection gradient is turned on (G_s). The phase-encoding gradients (G_p) change the phase of magnetization in the selected slice, and the frequency-encoding gradients (G_f) allocate specific frequencies. The echo is received while G_f is activated. B: "k-space"; Every echo generated from each phase gradient intensity, completes a line of the "k space" with the echo voltage values digitized. The "k space" thus generated contains as many lines as phase gradient intensities are included in the acquisition. C: Fourier transformation of the k space, originating the image in the frequency space, which we can interpret. (Adapted from [McRobbie et al., 2003, Wood and Wehrli, 1999])

FT of a continuous function $f(k)$ is

$$F(k) = \int_{-\infty}^{+\infty} f(x) e^{i2\pi kx} dx \tag{2.5}$$

where k represents the coordinates of the *time domain space* or *k-space*, and x its coordinates in the frequency domain. In the MRI experiment, this complex voltage signal is digitized, dismantled by the Fourier analyses, transforming the entire k-space into the image we can see and interpret. Figure 2.3, B and C panels, show an MRI image of a human head, as acquired instrumentally in the k space, and its final appearance after Fourier transformation to the frequency space, respectively. Large efforts were dedicated since the late 1970s to decrease the acquisition times in MRI techniques by improving k-space efficiency, without concomitant loss of signal properties or contrast. In that sense, revolutionary changes arrived with the improvement of technical acquisition strategies of (FSE) [Hennig et al., 1986] and echo-planar imaging (EPI) [Mansfield, 1984].

2.1.3 MRI sequences and contrast

The primary sources of endogenous tissue contrast in MRI are the hydrogen spin densities (N[H]), longitudinal relaxation times (T_1), and transverse relaxation times (T_2) [Hendrick and Raff, 1992]. Other sources, such as diffusion, perfusion and flow [Le Bihan et al., 1993], magnetization transfer [Van Buchem and Tofts, 2000], magnetic susceptibility inhomogeneity's [Pathak, 2009, Schenck, 1996] and

chemical shift [Brink et al., 1989, Brateman, 1986, Brady et al., 1985], affect also image contrast, but to a smaller extent. Contrast between two tissues (or voxels) A and B is expressed as $C_x = \frac{X_B - X_A}{X_B + X_A}$ where x represents the parameter evaluated by MRI. Most physiological properties of tissues (as diffusion and perfusion) will remain unchanged regardless of the imaging method implemented, while the inherent magnetic properties of tissues (T_1, T_2 and T_{2^*}) will depend on the on the external magnetic fields applied. Image contrast between two tissues (or pixels) may be defined analogously as the difference between the corresponding signal intensities.

In a spin echo sequence, the most commonly used pulse sequence in clinical MRI, the signal intensity S, depends on tissue parameters as follows

$$S(TE, TR) = N[H][1 - 2e^{-\frac{TR-TE/2}{T_2}} + e^{-\frac{TR}{T_1}}]e^{-\frac{TE}{T_2}} \tag{2.6}$$

By varying the TE and TR times, the SE sequence can be used to highlight T_1, T_2 or spin-density differences $N[H]$ between regions. Figure 2.4 provides examples of representative T_1 and T_2 weighted images of the same rat head bearing an implanted C6 tumor, as well as the mechanisms underlying T_1 and T_2 contrast as optimized using different TRs and TEs, respectively. In addition, many improvements have been implemented to decrease the long acquisitions required [Bradley et al., 1999] and sensitize the sequence to other variables including T_2^* and perfusion [Pathak, 2009, Chavhan et al., 2009], arterial and venous flow [Yiğit et al., 2017, Park et al., 2017], or even neuronal activation [Logothetis and Pfeuffer, 2004].

Figure 2.4, C and D panels, illustrate the origin of MR contrast in two tissues A and B with different T_1 and T_2 properties, using a spin-echo sequence. Figure 2.4 A depicts the MR signal intensity (%) of a SE sequence depending on the TR used, in two tissues (A and B) with different T_{1s}. Using a short-medium repetition time TR_1, the signal contrast obtained between the tissues A and B is the maximum. For longer times, like TR_2, longitudinal magnetization is almost recovered in both tissues and the signal contrast in T_1 is much lower. Figure 2.4 D shows the signal decay with increasing TE values, for a SE sequence in two tissues with different T_2 relaxation times. At short and medium TE's, like TE_1, the signal difference between tissues is very appreciable, while at long echo times, like TE_2, the signal received is almost zero for both tissues. In summary, optimizing TR and TE values allows to optimize T_1 and T_2 contrast.

2.1.4 Diffusion weighted MR imaging

Molecular diffusion refers to the random translational motion of molecules resulting from their thermal energy and is also known as Brownian motion. In a free medium, and during a given time interval, molecular displacements follow a Gaus-

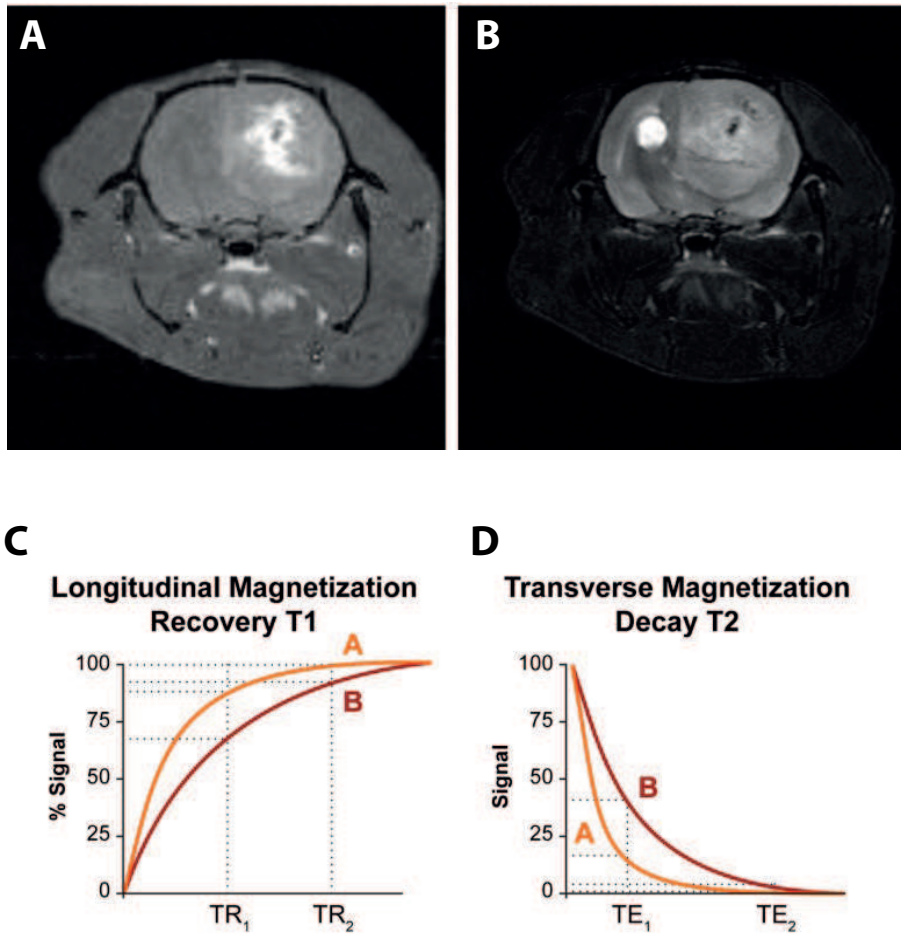

Figure 2.4: Representative spin-echo images from a rat brain containing an implanted C6 glioma tumor. Different weightings are obtained by varying TR and TE (Equation 2.7). A: T_{1w} with GdDTPA (TR=300ms, TE=10ms), B: T_{2w} (TR=3000ms, TE=60ms). C: Signal intensities from two tissues with different T_1, as function of TR. D: Signal values of two tissues with different T_2 relaxation times in terms of the TE used. Values are expressed in % of the signal corresponding to the total magnetization recovery. (Modified from [Hendrick and Raff, 1992]).

sian distribution and travel randomly over a distance, well described by a diffusion coefficient (D). This coefficient depends on the mass of the molecules, the temperature and the viscosity of the medium. A water molecule diffusing in water at 37 °C, has a diffusion coefficient of $3 \cdot 10^{-9} m^2/s$, which means that travels, on average, 17 μm during 50 ms (Figure 2.5 A). During their diffusion-driven displacements, molecules probe tissue structure on a microscopic scale; they cross or interact with cell membranes, macromolecules or fibers, that modify the water molecule diffusion properties, reducing the expected diffusion time and definitely altering the Gaussian distribution (Figure 2.5 B). On these grounds, the non-invasive quantification of water diffusion distributions *in vivo* provides unique evidences on structural and geometric characteristics of tissues. Moreover, diffusion measurements are also perfectly endowed to probe potential changes in physiological or pathological states [Le Bihan, 2003].

The MRI methodology used to evaluate diffusion of water molecules in biological tissues is known as Diffusion Weighted Imaging (DWI). It was first proposed by Le Bihan and co-workers [Le Bihan et al., 1986] and has emerged, since then, as one of the most utilized MRI protocols in clinic [Sotak, 2002, Mayer et al., 2010]. The sequence is based on the Stejskal-Tanner (ST) method [Stejskal, 1965], which introduces a couple of magnetic field gradients at both sides of the 180°pulse in a SE technique (Figure 2.5 A). Those gradients are capable of encoding water molecules in such a way that the final echo received will be lower if water molecules have diffused away from the slice, in-between the application of the phase gradients, or higher if those molecules have not diffused from the slice –or have moved in a coherent manner.

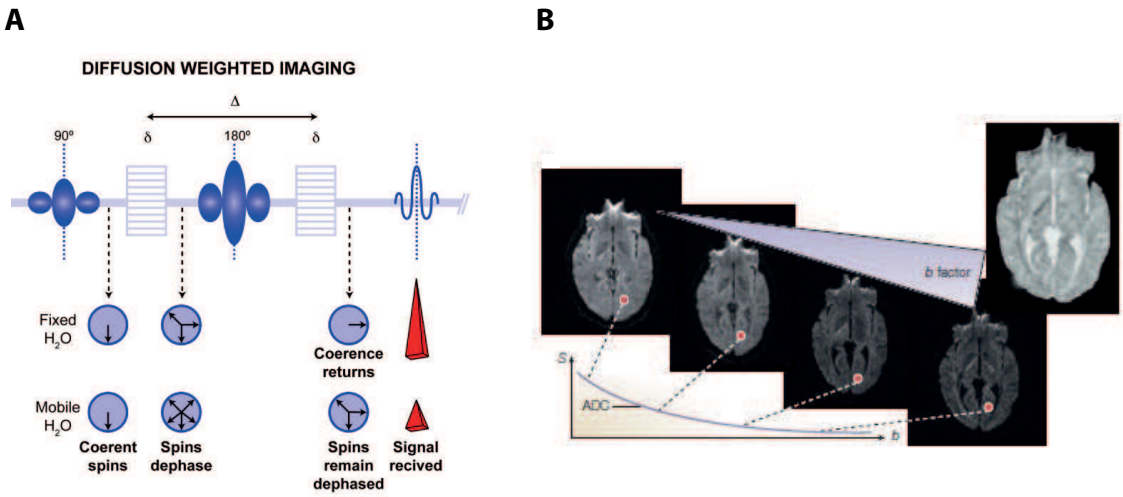

Figure 2.5: Fundamentals of Diffusion Weighted Imaging (DWI). A: Schematic representation of a water diffusion encoding MRI sequence. After the 90º RF pulse, a magnetic field gradient is applied during a specific time (δ) and spins in water molecules experience a specific dephasing. After a time ($\Delta - \delta/3$), a second, and exactly opposite gradient is applied behind de 180º pulse. If molecules in a voxel have moved incoherently, spins will not be able to recover the same phase and the total magnetization in a voxel will loss coherence, hence intensity. Conversely, if motion is coherent, all spins will experience the same rephasing and final coherence will not change. B: Diffussion encoded decrease in MRI signal intensity for increasing b factor values. The Apparent Diffusion Coefficient may be fitted from this exponential and represented as an ADC map (Modified from [Le Bihan, 2003]).

The sensitivity of a DWI acquisition depends on the gradient duration (δ), gradient separation (Δ) and gradient strength (G). The imaging protocol typically consists in applying a series of magnetic field gradients with different diffusion sensitivity, and fitting the MRI signal obtained to an appropriate diffusion model of water in biological tissues (Figure 2.5 B).

Several models of water diffusion within biological tissues have been proposed [Szafer et al., 1995], including mainly the mono-exponential model [Syková and Nicholson, 2008], the bi-exponential model [Lizarbe et al., 2013, Lizarbe et al., 2015] or statistical approaches as q-ball imaging [Jensen et al., 2017, Caiazzo et al.,

2016, Jensen et al., 2016, Tuch, 2004] or kurtosis imaging [Huang et al., 2015].

The mono-exponential model assumes that the diffusion of water results in a monoexponential decay of the diffusion weighted signal given by the expression

$$\frac{S(\delta, \Delta, G)}{S(0)} = e^{-\gamma^2 \delta^2 G^2 (\Delta - \delta/3) ADC} \tag{2.7}$$

where $S(\delta, \Delta, G)$ represents the intensity of diffusion weighted signal in the presence of a G gradient of duration δ, during a diffusion interval Δ, S(0) the intensity of the diffusion signal in the absence of diffusion gradients and ADC the apparent diffusion coefficient of water. The term $-\gamma^2 \delta^2 G^2 (\Delta - \delta/3)$ is normally grouped under a single value termed b, in recognition for Denis Le Bihan, who first proposed its use.

The biexponential model proposes that water diffusion in tissues may be best described by a biexponential function, including physiological environments of "fast" and "slow" water diffusion, as indicated by the expression

$$\frac{S(b)}{S(0)} = SDP \cdot e^{-bD_{\text{slow}}} + FDP \cdot e^{-bD_{\text{fast}}} \tag{2.8}$$

where S (b) and S(0) represent the signal intensities in the presence and absence of diffusion sensitizing gradients, D_{slow} and D_{fast} the apparent diffusion coefficients in the "slow" and "fast" diffusion environments and SDP and FDP the corresponding proportions of slow and fast diffusion phases.

The Q-ball method, proposed by David S. Tuch [Tuch, 2004], used in tractography, is an implementation of the high angular resolution diffusion imaging (HARDI), able to resolve some limitations of the tensor model, notably the one hampering the resolution of more than one tract occurring per voxel examined [Jensen et al., 2017, Caiazzo et al., 2016, Jensen et al., 2016]. However, the method requires large pulsed field gradients and time-intensive sampling. Recently, it has been shown that the HARDI signal can be reconstructed model-independently using a spherical tomographic inversion called the Funk-Radon transform, also known as the spherical Radon transform. The resulting imaging method, termed q-ball imaging, can resolve multiple intravoxel fiber orientations and does not require any assumptions on the diffusion process such as Gaussianity or multi-Gaussianity.

While the monoexponential and bi-exponential models assume Gaussian diffusion of water, several evidences suggest that water diffusion within tissues is non-gaussian, displaying some degree or kurtosis [Steven et al., 2014]. Imaging this kurtosis may become an important tool to improve our understanding of the diffusion processes in biological tissues and extend it to the clinic [Jensen et al., 2017, Hori et al., 2012]. Briefly, tissue diffusion is best described by the expression

$$Ln(S(b)/S(0)) = -b \cdot ADC + 1/6b^2 \cdot ADC^2 \cdot K_{\text{app}} \tag{2.9}$$

where K_{app} represents the apparent kurtosis along a defined diffusion direction. The kurtosis metrics and corresponding kurtosis imaging are able to reveal structural features not accessible to conventional DWI, at the expense of significantly longer acquisition times and number of diffusion directions explored.

In general, all diffusion models depict advantages and disadvantages; the most important advantages being to explore tissue microstructure in vivo, and to investigate the physiological mechanisms underlying DWI images. The more important disadvantages rely on the assumptions of specific diffusional behaviors or the use of simplified phenomenological models of water diffusion. To overcome this fundamental limitation, in particular, I found it useful to implement in this thesis model-independent interpretations of diffusion weighted images .

2.1.5 Applications

The applications of MRI, initially restricted to few specialized laboratories, have reached widespread diffusion in preclinical and clinical medicine. Both structural and functional MRI approaches have been implemented to characterize the mechanisms underlying physiological function, or the imaging signatures of disease. In general, most physiopathological processes haven been imaged, emphasizing those affecting the brain, with special reference to cerebral activation, ischemic episodes or neurodegenerative disorders. In the following sections, I will describe briefly the most relevant physiopathological characteristics, and MRI approaches, used to investigate obesity and cancer, the two diseases approached in this thesis.

Appetite control

Obesity is a pandemic syndrome, underlying the most abundant, morbid and prevalent diseases in developed countries, including atherosclerosis, ischemic episodes, neurodegeneration and cancer [Kopelman, 2000, Kahn et al., 2006]. The obesity syndrome is caused by an imbalance in appetite regulation resulting in a positive balance between food intake and energy expenditure [Coll et al., 2007, Coll and Yeo, 2013, Morton et al., 2006]. Its effective treatment remains unresolved today, representing a daunting scientific and medical challenge, with already enormous socio-economical repercussions [Blair and Nichaman, 2002].

Considerable progress has been achieved in the last decades on the molecular and physiological mechanisms underlying appetite regulation. In spite of this, the direct evaluation of appetite regulation in the brain remains difficult, very often limited to indirect measurements of food intake and body weight.

On these grounds, the implementation of non-invasive methodologies for the

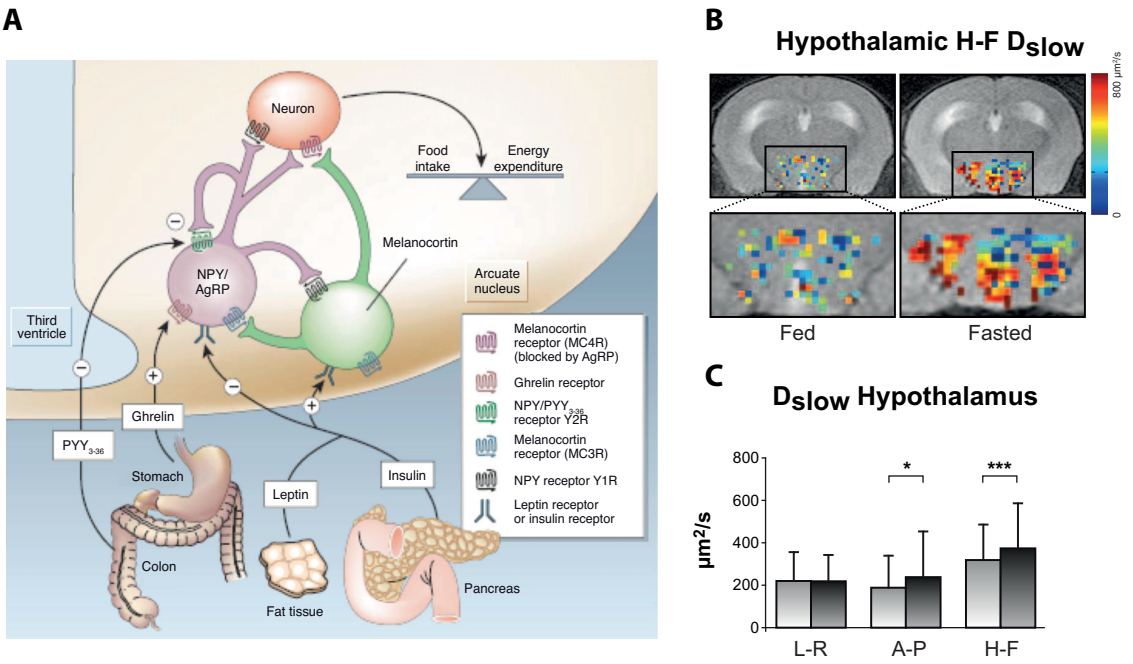

Figure 2.6: Systemic and intrahypothalamic mechanisms of appetite regulation. A: Neuronal mechanisms regulating appetite by peripheral organs and associated hormonal signals and intrahypothalamic neuropeptides produced by orexigenic and anorexigenic neurons (Taken from [Coll et al., 2007]). B: Representative DWI images from the hypothalamus of a fed and a fasted rat, obtained using the biexponential model (Taken from ref [Lizarbe et al., 2013]).

evaluation of appetite entails considerable relevance. A variety of neuroimaging tools have been proposed including mainly positron emission tomography (PET) and functional Magnetic Resonance Imaging (fMRI) methods [Carnell et al., 2012]. PET studies provide information on cerebral activation by detecting the emitted positrons derived from the increased uptake of ^{18}F-deoxiglucose, an event revealing the metabolic coupling between glucose uptake, blood flow and neuronal activity in the hypothalamus during feeding-related stimuli [Gautier et al., 2000, Tataranni et al., 1999]. Blood Oxygenation Level Dependent (BOLD) fMRI, infers regional neuronal activity from the changes in magnetic susceptibility that take place during the transition between oxygenated and deoxygenated haemoglobin, occurring after increased oxygen delivery to activated neurons in the hypothalamus of rats and humans [Mahankali et al., 2000, Matsuda et al., 1999]. Finally, Manganese Enhanced Magnetic Resonance Imaging (MEMRI), uses manganese ion accumulation as a surrogate marker of the increased calcium movements occurring during neuronal activation [Parkinson et al., 2009]. Notably, these previous approaches are not devoid of limitations, mainly derived from their reduced spatial and temporal resolution in the PET and BOLD fMRI approaches, and the potential neurotoxicity of Mn^{2+}, in the MEMRI technique. Moreover, these earlier studies showed that appetite stimulations occurred within a narrow dynamic range over much larger resting values, involving the use of critical image subtractions and statistical parametric mapping strategies [Ashburner, 2012, Ashburner, 2009], an approach recently

shown to require further validation in some cases [Eklund et al., 2016].

To overcome these limitations, we proposed the use of functional Diffusion Weighted Imaging fDWI [Le Bihan, 2003], a novel functional technique improving the spatial and temporal resolution provided earlier by PET or BOLD and avoiding the use of the potentially toxic doses of manganese precluding its use of MEMRI in humans. However, previous fDWI interpretations relied on the use of phenomenological models of diffusion, an approach involving considerable uncertainties on the compartmental origin of the signals analyzed. Moreover, previous approaches considered only hypothalamic activation, neglecting appetite activation in other cerebral regions, including mainly the "reward" or "hedonic" centers (Figure 2.6). On these grounds, I found it useful for my PhD thesis, to propose an automatic approach to obtain, unambiguously, integral maps of the appetite impulse through the brain of rodents and humans.

Brain tumor diagnosis and treatment

Cerebral tumors originate from genetic alterations in neural stem cells resulting in uncontrolled progression and growth, development of neovasculature, and metabolic reprogramming, as required to support the exacerbated tumoral growth [De-Berardinis et al., 2008, Ward and Thompson, 2012] (Figure 2.7). Briefly a common progenitor cell is the ancestor of all brain cells, differentiating to neuronal, astroglial or oligodendroglial precursor cells [Singh et al., 2004]. The neuronal, astroglial or oliogodendroglial precursor may eventually, differentiate erroneously, to tumor cells originating medulloblastomas, gliomas or oligodendrogliomas [De La Rocha et al., 2014]. Gliomas represent the most frequent brain tumors, classified in grades I-IV of increasing malignancy [Verhaak, 2016]. In order to prolong the global survival rate, surgical treatments, radiotherapy and chemotherapy, have been implemented either alone or in combination, providing up to date, the most recommended therapeutic approaches. However, despite important progress in these treatments, gliomas remain currently among the highest mortality tumors.

Neuroimaging methods have played an important role in the diagnosis of gliomas. PET ^{18}FDG methods have been able to localize glioma neoplasms because of their inherently high glucose uptake, or increased ^{11}C choline metabolism, among other biomarkers [Hutterer et al., 2014, Puttick et al., 2015]. MRI methods have been proven successful in localizing the anatomical location of the lesions and providing vital pre-surgical information [Hutterer et al., 2014, Puttick et al., 2015].

Briefly, three generic approaches are used in glioma therapy; surgical removal, chemotherapy and radiotherapy. Neurosurgical removal remains the most effective treatment, if the anatomical location of the tumor allows safe resection. Chemotherapy approaches use normally, either DNA mitogens as temozolamide or irinotecan [Taal et al., 2015, Nicholas et al., 2005, Neyns et al., 2005] or antiangiogenic ap-

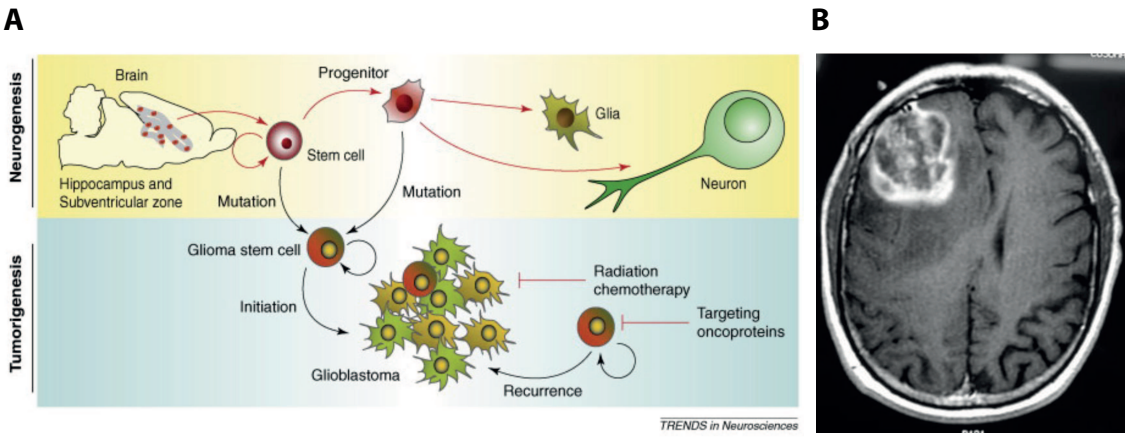

Figure 2.7: Hypothesis of glioma development from neural stem cells and representative MR image. A: Brain tumors originate from abnormal transformations of the common neuronal or glial progenitor cells. Primary tumors derive from direct transformation of the progenitor cells, while secondary tumors originate by subsequent transformations of previously transformed neurinoma, oligodendroglioma o astrocytoma cells (taken from [De La Rocha et al., 2014]). B: Representative T_{1W} MRI of a glioblastoma in a human brain. Note that contrast agent uptake occurs mainly in the growing rim of the tumor surrounding the necrotic area (taken from [Sorensen et al., 2008]).

proaches as bevazizumab, to interfere with the Vascular Entotelial Growth factor receptor [Khasraw et al., 2014, Cea et al., 2012]. Radiotherapy uses highly focused y radiations to induce DNA damage and necrosis, only in the tumoral region [Ryken et al., 2015, Brada, 2006]. Since a correct personalized diagnosis determines the most adequate treatment, there is an urgent need to improve personalized diagnostic and prognostic indicators to increase the efficacy of treatments and survival rates at the individual level. In this thesis, we will investigate automatic methods to detect the glioma lesions, how they develop and how they respond to chemotherapeutic or radiotherapy treatments.

2.2 Image pre-processing

The MR images obtained as indicated above, are influenced by a variety of factors including; intensity and duration of the electromagnetic irradiations used, magnetic loading of the receiver and transmitter coils, acquisition time, stability of the static magnetic field, gradient stability and intensity, the examined subject itself, etc. These factors make images to contain one or more unwanted contributions including; random noise, artifacts, contrast interferences, deformations, etc. difficulting their direct analysis and interpretation. In order to remove noise, redundant or irrelevant information from the acquired data, it becomes necessary to apply one or more image pre-processing steps, prior to the image analysis method. In the following sections, I will describe in general terms, a collection of steps that are often used during

the pre-processing of MR images from brain, including mainly four categories: slice selection and brain extraction, bias field correction, intensity normalization, and registration (Figure 2.8).

A variety of methods exist that can be used either individually, or chained as a pipeline, during the pre-processing phase, all of them with different requirements and objectives. The performance of each one of these methods has been generally tested individually in the literature using different testing environments and evaluation metrics. The results show that, based on the existing literature, it is difficult to reach general conclusions about the optimal combinations of pre-processing algorithms. Relevant manuscripts address this problem [Strother, 2006, Churchill et al., 2012a, Churchill et al., 2012b, Glasser et al., 2013, Churchill et al., 2015, Churchill et al., 2017, Vergara et al., 2017], proposing a framework to evaluate different pipelines for the same dataset or optimizing different pipelines for more specific cases. Despite the favorable results achieved in these studies, choosing the optimal pre-processing method remains as an open problem of current medical image analysis.

2.2.1 Slice selection and brain extraction

MR image acquisitions usually contain many slices from the investigated region of interest in a subject. However, the anatomical structure investigated, may be present in only one, or just a few, of the acquired slices. It becomes, therefore, necessary to begin the analysis by selecting, among those acquired, only the relevant slices to be analyzed (Figure 2.8 B). Generally, this is a non-automatic procedure requiring considerable experience in the slice(s) selection process.

Moreover, even though the adequate slice may have been selected, MR images of brain usually contain non-brain structures like muscle, skull, skin or fat, that may interfere with the automatic analysis of the cerebral tissue contained in these images. Therefore, an important operation is required to isolate only the brain, from surrounding non-cerebral tissues. This operation is referred to as skull stripping or brain extraction [Hata et al., 2005, Fennema-Notestine et al., 2006, Mahmood et al., 2015, Chaddad and Tanougast, 2016]. Several methods of human brain extraction in MRI have been proposed, including: Brain extraction tool (BET) [Smith, 2002], BET2 [Jenkinson et al., 2005], Brain surface extraction (BSE) [Shattuck et al., 2001], and Hybrid watershed algorithm (HWA) [Ségonne et al., 2004]. BET is designed for T_1-weighted images and uses a set of locally adaptive model forces and thresholding to extract the brain contour. One limitation of BET is that it fails to extract the brain region in the bottom of axial slices, since in these cases, the head scan may include the neck portion, and the center of gravity of the volume lies outside the brain, the deformation including inevitably non-brain portions [Kalavathi and Prasath, 2016]. BET2 is based on BET and uses techniques of intensity clamping, surface point detection, and mesh fitting to extract the brain. This method needs

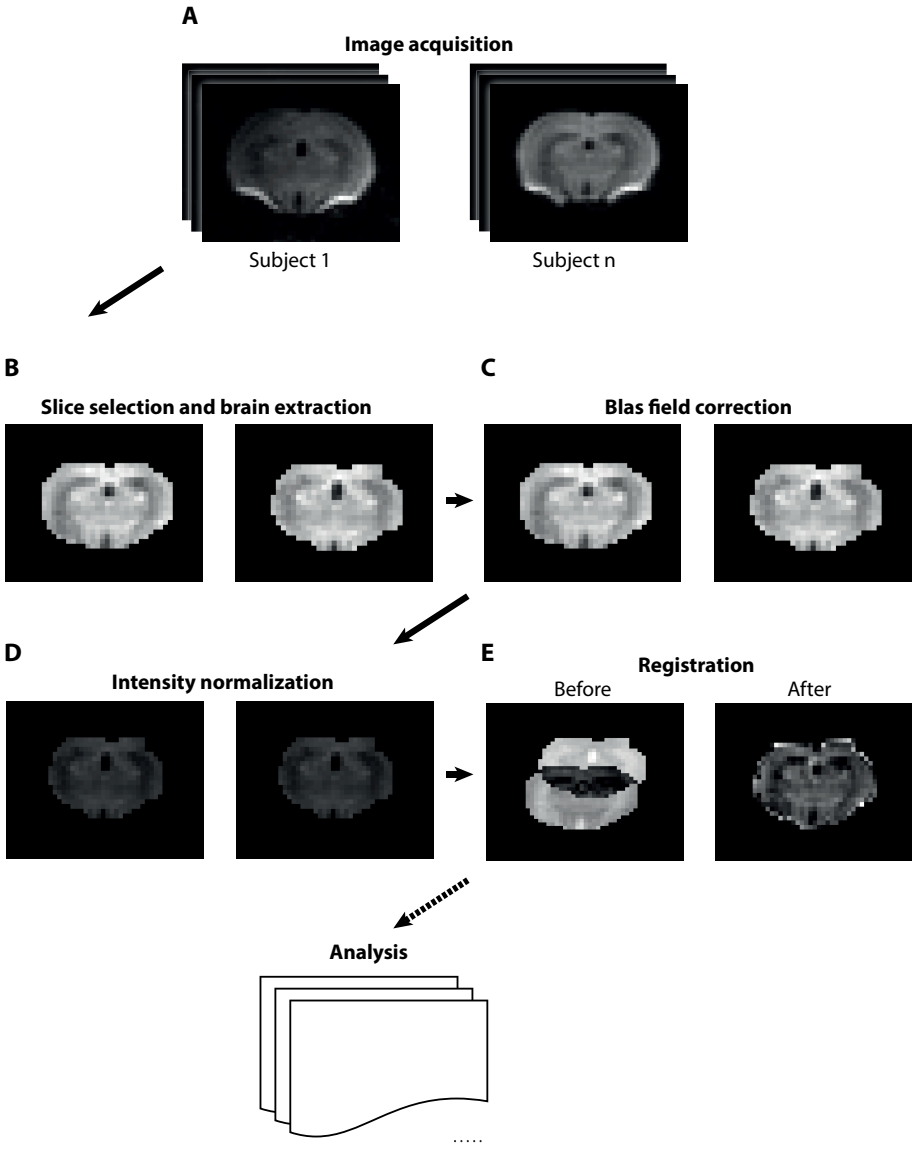

Figure 2.8: Flow diagram with the most frequently used pre-processing steps. A: Image acquisition. B: Slice selection and brain extraction. C: Bias field correction. D: Intensity normalization. E: Registration.

to receive both T_1 and T_2 images as input, which may constitute a limitation in situations where these images have not been acquired. BSE conversely, can be applied on T_1 and T_2-weighted images to extract the brain, mainly using techniques based on anisotropic diffusion filtering and edge detection. This method has the limitation that sometimes dura mater may also be included in the brain mask and therefore the edge detector may not find a clear brain boundary [Kalavathi and Prasath, 2016].

HWA needs to receive T_1-weighted images as input and combines watershed algorithm [Hahn and Peitgen, 2000] and deformable surface model [Dale et al., 1999] to extract the brain. A relevant limitation of this method is that it may include the non-brain portion of the segmented brain due to intensity bias in the input image [Kalavathi and Prasath, 2016]. HWA has been shown to have the highest sensitivity in general but the lowest specificity in a comparative study on various methods [Fennema-Notestine et al., 2006, Hartley et al., 2006, Shattuck et al., 2009, Kalavathi and Prasath, 2016]. Most of the existing skull stripping methods are applicable to T_1-weighted MR brain images. Interestingly, HWA appears to be more robust to the change of parameters than the other methods. BSE had higher specificity, and BET always under-segmented, including additional non-cerebral tissues. Notably, none of these methods provided satisfactory performance when evaluated with a wide range of scan types or orientations, probably because of the complexity and variations in the human brain structures, presence of image noise, inadequate image contrast, or image artifacts [Somasundaram and Kalavathi, 2011, Somasundaram and Kalavathi, 2012, Kalavathi and Prasath, 2016].

Several methods have been proposed for rodent brain extraction specifically, since those implemented for humans often perform poorly in rodents because of the inherent differences in shapes and contrasts [Chang et al., 2017]. Among the rodent brain extraction methods in MR images, the following merit specific mention: 2-D [Murugavel and Sullivan, 2009] and 3-D [Chou et al., 2011] Pulse-Coupled Neural Network (PCNN), a method based on BET [Li et al., 2013] and based on gray-scale mathematical morphology and LOGISMOS-based graph segmentation [Oguz et al., 2014]. More recently, High-order and L0 regularized Mumford-Shah model (HoL0MS) was proposed for rodent and human brain MR images, its main limitation relying in the processing the T_1-weighted MR images because of the low contrast between the skull and the cerebrospinal fluid [Chang et al., 2017]. An example of slice selection and brain extraction using PCNN algorithm is shown in Figure 2.8 B.

2.2.2 Bias field correction

Spurious smoothly varying image intensities, shading, or bias field may be dealt specifically with Bias field correction pre-processing methods. Intensity inhomogeneities in MRI, arise from imperfections in the image acquisition processes, appearing

normally as a smooth intensity variation across the image (Figure 2.8 A, subject 1). This phenomenon causes the intensity of the same tissue to vary with its location within the image. These inhomogeneities are normally hardly noticeable to a human observer, but many medical image analysis procedures, such as segmentation and registration, become indeed highly sensitive to these spurious variations in image intensities [Vovk et al., 2007].

Several methods to correct intensity inhomogeneities in MRI have been proposed and validated [Belaroussi et al., 2006, Hou, 2006, Vovk et al., 2007]. Figure 2.8 C, subject 1 illustrates the results of bias field correction using a polynomial regression of second order.

Non-parametric non-uniform normalization (N3) [Sled et al., 1998] is the most popular method to solve this problem. N3 is completely automatic, does not require prior knowledge, and can be applied to virtually all types of MRI. This method is iterative, seeking the smooth multiplicative field that maximizes the high frequency content of the distribution of tissue intensity, using a B-spline smoothing strategy [Vovk et al., 2007]. In addition, N3 is public and open source. Several comparative studies have shown the excellent performance of N3 and depicted that certain parameters associated with the B-spline least-squares fit are essential for an adequate performance [Arnold et al., 2001, Boyes et al., 2008, Zheng et al., 2009].

More recently, an N3 variant approach has been reported, N4ITK [Tustison et al., 2010]. Compared to N3, N4ITK includes a faster and more robust B-spline approximation routine with a modified hierarchical optimization scheme, improving the intensity inhomogeneity correction. Notably, N4ITK is also open source and publicly available.

2.2.3 Intensity normalization

Pixel intensity range values can change from one image to another within the same MRI protocol, for the same subject and body region, within images obtained from the same patient and scanner in different moments, or even from images of the same patient obtained in different scanners. Intensity range variations may be caused by inevitable differences in image acquisition conditions including; different coil loadings, static field inhomogeneity's and patient positioning, temperature fluctuations, or random signal noise from external (or internal) sources entering the detection coils. These variations interfere significantly with important post-processing tasks such as automatic segmentation or classification routines, which depend strictly on the observed image intensities, thus compromising the accuracy and complicating the interpretation [De Nunzio et al., 2015]. Standardization routines minimize significantly the variation of the overall mean intensity of the MR images [Nyúl et al., 1999, Nyúl et al., 2000, Ge et al., 2000] obtained from the same tissue region in

studies obtained on the same or different scanners [Madabhushi and Udupa, 2006].

Nyul and Udupa presented a popular intensity normalization method based on nonlinear transforms [Nyúl et al., 1999]. This is a twostep process wherein all images are transformed in such a way that, for the same protocol and body region, similar intensities will have a similar tissue-specific significance. Firstly, the parameters of the standardizing transformation are learned from a set of images. Secondly, for each MRI study, these parameters are used to map the original intensity gray scale into the new, normalized, gray scale. The mode on the histogram is used normally as the landmark for transforming the scene intensities. Nyul and Udupa [Nyúl et al., 2000] showed later that the mode was not a sufficiently robust landmark, describing a variant of the original standardization protocol, replacing the mode by the median and other quartile locations on the histogram. These improvements were shown to become more robust than the original mode-based method [Madabhushi and Udupa, 2006].

In [Madabhushi and Udupa, 2006], authors explored the main weakness of these approaches. For example, sometimes the same histogram-based landmark may represent one tissue of a patient, while in other patients it may represent different tissues, a circumstance occurring frequently in studies in which significant changes in image intensity among patients happen. Along these lines, Madabhushi and Udupa proposed two new more robust intensity normalization methods than landmarks derived from image intensity histograms. The methods exploited the ability of the g and gB-scale to automatically partition the image into homogeneous regions, corresponding to different tissue regions. To perform the nonlinear mapping of intensities, the new scale-based methods use, unlike the preexisting methods, landmarks derived from the individual scale regions of the image.

A different approach was implemented in [Cao et al., 2016, Lizarbe et al., 2015, Galla et al., 2017], where mono and bi-exponential models were fitted to diffusion weighted MR images. The normalization applied consisted in dividing each pixel intensity value by the intensity value of the same pixel in basal, non-diffusion weighted, image (Figure 2.8 D).

2.2.4 Registration

In medical studies aimed to compare several images, previous transformations are usually required because images may have been acquired with different imaging modalities, scales, orientations, acquisition conditions or anatomical coordinates. These transformations are applied in a pre-processing step called registration (Figure 2.8 E), which has reached fundamental importance for the correct analysis and interpretation of biomedical images.

The registration setup can be classified [Maintz and Viergever, 1998, Viergever et al., 2016] according to: dimensionality (spatial or spatio-temporal 2D/2D, 2D/3D, 3D/3D), nature of the registration base (extrinsic, intrinsic, non-image based), nature of the transformation (rigid, affine, projective, curved), domain of the transformation (global, local), degree of interaction (interactive, semi-automatic, automatic), optimization procedure (parameters computed or searched for), modalities involved (mono-modality, multi-modality, modality to model, patient to modality), subjects involved (intra-subject, inter-subject, atlas), or organs involved (e.g., brain, heart, breast).

Recent work has focused on improving techniques of medical image registration to find new methods, or extend existing ones, to obtain greater accuracy and robustness in the alignment [Hill et al., 2001, Keszei et al., 2017, Oliveira and Tavares, 2014, Slomka and Baum, 2009]. Considerable efforts have been devoted to curved transformations and inter-subject registration methods, yielding very valuable results [Keszei et al., 2017, Klein et al., 2009].

In particular, Klein et al. evaluated 14 nonlinear deformation algorithms applied to human brain MRI registration, concluding that Symmetric Normalization (SyN) [Avants et al., 2008] and ART [Ardekani et al., 2005] provided consistently high ranking results and were the only methods that reached top rank in all tests and for all label sets [Klein et al., 2009].

Avants et al. [Avants et al., 2011] presented later an open source tool software package called Advanced Normalization Tools (ANTS), containing a variety of transformations ranging from rigid, affine, and elastic to diffeomorphic models, including the SyN transformation.

However, it should be noted here that despite the important advances made in the development of good alignment algorithms, it is still necessary to improve the following two aspects: (i) validation of registration methods and (ii) translation of results of image registration research into clinical practice [Viergever et al., 2016].

2.3 MRI analysis: conventional approach

In research environments, image studies are frequently based in a relatively small number of subjects. In these cases, it is very common that researchers implement a detailed manual analysis of the images. For example, in order validate a study hypothesis, a region of interest (ROI) is manually selected on the acquired images and subsequently the analysis focuses on that particular region. After the ROI is selected, a mathematical model is chosen to fit the distribution of the pixel intensities and a parameter is computed from it. Finally, a statistical test can be applied assessing the validity of the study hypothesis. Many publications in the field of

medical image analysis have used this approach [Cao et al., 2016, Lizarbe et al., 2013, Lizarbe et al., 2015, Galla et al., 2017].

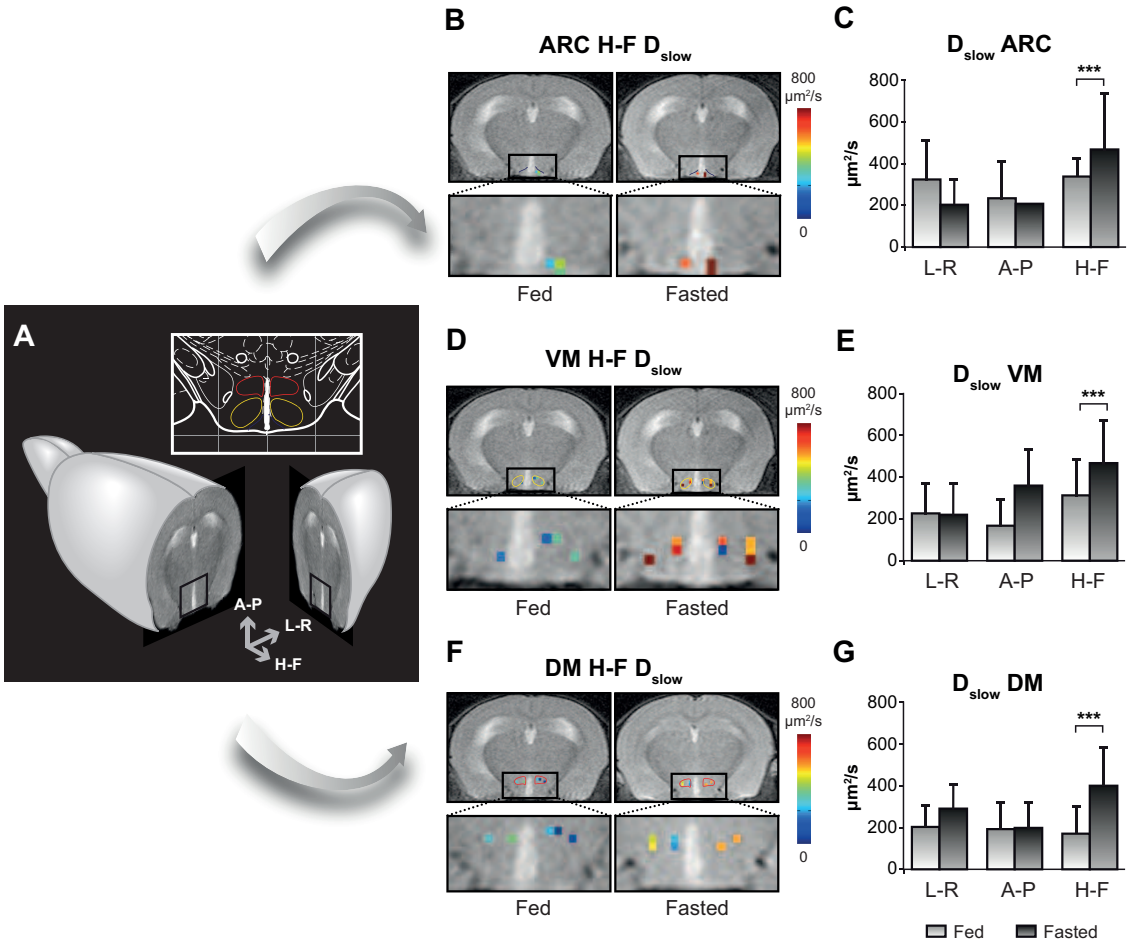

Figure 2.9: Representative user-based ROI analysis. Imaging appetite by fDWI in the hypothalamic nuclei of the mouse brain, obtained using the biexponential model (Taken from [Lizarbe et al., 2013]).

Lizarbe et al. (Figure 2.9) illustrated this approach with a non-invasive visualization of hypothalamic activation by fasting in the mouse and human brains, using diffusion weighted magnetic resonance imaging and the bi-exponential model. Their results showed that hypothalamic activation by fasting can be detected non invasively by diffusion weighted magnetic resonance imaging through changes in the water diffusion parameters both in mice and humans [Lizarbe et al., 2013]. More recently, Lizarbe et al. extended this approach to evaluate the contribution of leptin-dependent anorexigenic pathways and neuropeptide Y (NPY)-dependent orexigenic pathways to the changes in hypothalamic water diffusion parameters observed in vivo by functional diffusion weighted MRI [Lizarbe et al., 2015].

Cao et al. investigated whether DWI, with apparent diffusion coefficient (ADC) measurements, can detect skull bone marrow infiltration in newly diagnosed acute

lymphoblastic leukemia (ALL) children, before therapy and normalization in complete remission after treatment. The decrease in mean ADC value of infiltrated tissue revealed diffusion restrictions due to hypercellularity of the tumoral tissues. When complete clinical remission was achieved following successful treatment, DWI signal abnormalities, as well as mean ADC values, returned to normal. Their study suggested DWI as a new tool for monitoring oncologic treatments in ALL [Cao et al., 2016].

Finally, Galla et al. assessed the use of apparent diffusion coefficient (ADC) in predicting length of survival after superselective intra-arterial cerebral infusion (SIACI) of bevacizumab (BV) and overall survival in patients with recurrent glioblastoma. In patients with recurrent glioblastoma treated with SIACI BV, the change in ADC value after treatment was found to be predictive of overall survival [Galla et al., 2017].

2.3.1 Standard software packages

Several statistical software packages have become widely available and routinely used in image analysis recently as, Statistical Parametric Mapping (SPM) [Ashburner, 2012, Penny et al., 2011], FMRIB Software Library (FSL) [Jenkinson et al., 2012], and Analysis of Functional NeuroImages (AFNI) [Cox, 1996, Cox, 2012]. These packages contain pre-processing and analyzing tools, and can be used to analyze the effects of sensorial or motor stimulation through the whole brain. They examine the images voxel by voxel, looking for the existence of stimulus related differences, substantiated through statistical tests.

SPM is freely distributed software designed for constructing parametric statistical maps for the analysis of brain imaging data sequences [Penny et al., 2011]. The sequences can be a series of images of different planes and conditions, or temporal series of the same subject. The current version is implemented for the analysis of images of different nature: functional magnetic resonance, positron emission tomography, single photon emission tomography, electroencephalography and magnetoencephalography.

Figure 2.10 provides a visual summary of the image processing steps in SPM. First, image time-series are pre-processed applying registration, smoothing, and normalization methods. Next, numerous hypothesis tests can be performed, assuming that each voxel can be modeled using the General Linear Model (GLM) [Friston et al., 1994]. The result of the statistical analysis is a p-value for each voxel, which represents the probability of the absence of significant effects. The result is an image called “statistical parametric map”, whose voxel values are p-values. SPM integrates the theory of GLM and random field theory (RFT) [Worsley et al., 1992] to perform classical inference on topological features of statistical parametric maps.

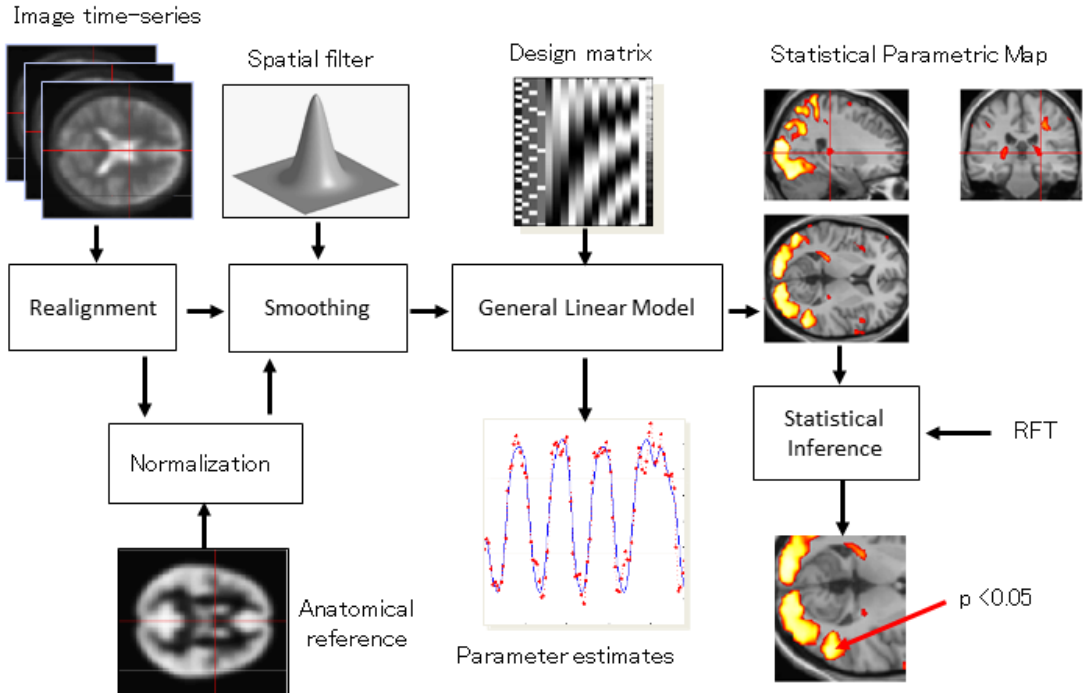

Figure 2.10: Visual summary of the SPM approach. (Taken from [Nair and Halahakoon, 2016]). SPM approach can be summarized in three stages: images pre-processing (left), application of the General Linear Model (center) and results in form of statistical parametric map using the random field theory (right).

FSL is another open source library of analysis tools for functional, structural and diffusion MRI brain imaging data. In FSL, there is a set of major tools that can be used directly (Table 2.1).

Finally, AFNI is an open source software package for the analysis and visualization of three-dimensional human brain functional MRI data. The tools included in AFNI can be used in a flexible manner, interactively as isolated tools or chained as a complete pipeline with pre-processing and analysis steps (Table 2.2).

Even though these packages are mainly dedicated to process human brain images, SPM has a specific toolbox, SPMouse, which extends SPM functionality with affine registration priors for the mouse and marmoset brains [Sawiak et al., 2009].

Definitely, SPM, FSL, and AFNI are powerful tools to analyze brain images but their efficient and correct use requires deep knowledge of the important underlying statistical and biological theories. These packages assume that data can be modeled using GLM where noise is additive and independent, noise is Gaussian with zero mean and variances are equal among groups [Penny et al., 2011]. In practice, this assumption is rarely strictly met, although the method is relatively robust given slight deviations from this hypothesis [Petersson et al., 1999].

Tool name	Use
FEAT	FMRI: task-based, using GLM
MELODIC	FMRI: resting-state or task-based, using ICA and no temporal model
FABBER	ASL (perfusion imaging of flow)
FDT	Diffusion: probabilistic tractography
TBSS	Diffusion: multisubject voxelwise analysis
BET	Brain extraction
FAST	Tissue-type segmentation (GM/WM/CSF)
FIRST	Subcortical segmentation
FLIRT and FNIRT	Linear and Non-Linear Registration
FSL-VBM	Voxel-wise analysis of grey matter density
SIENA and SIENAX	Whole brain atrophy (longitudinal and cross-sectional)

Table 2.1: Major tools in FSL. (Adapted from [Jenkinson et al., 2012])

Functional MRI is now twenty five years old, yet its most common statistical methods have not been validated using real data until recently [Eklund et al., 2012, Eklund et al., 2016]. The work of Eklund et al. [Eklund et al., 2016], is particularly meritorious in this respect. Briefly, these authors used real resting-state data and analyzed a total of 3 million random task groups to compute empirical familywise error rates for the software packages SPM, FSL, and AFNI, as well as a nonparametric permutation method. Notably, for a nominal familywise error rate of 5%, the parametric statistical methods were shown to be conservative for voxelwise inference but invalid for clusterwise inference. Results from this work suggested that the principal cause of invalid cluster inferences relies in spatial autocorrelation functions that do not follow the assumed Gaussian shape. By comparison, the nonparametric permutation test is found to produce nominal results for voxelwise as well as clusterwise inference. Authors commented that these findings demonstrate the need to validate the classical statistical methods being used in the field of neuroimaging.

2.4 MRI analysis: machine learning approach

"Machine learning refers the development of computer algorithms able to unravel complex relationships or patterns from empirical data and make accurate decisions" [Bishop, 2006, Duda et al., 2012, Mitchell, 1997, Wang and Summers, 2012]. As an interdisciplinary field, it is related to statistics and probability theory, in order to define and understand the proposed models, and to theoretical computer science and optimization in order to train these models. Machine learning is currently being exploited in a wide variety of real life problems such as medical diagnosis, self-driving cars, product recommendations, publicity campaigns and fraud detection.

Tool name	Use
to3d	Conversion of 2D image files into 3D datasets for use with AFNI
3dvolreg	Slice timing correction and motion correction of functional images
3dDeconvolve	Single-subject analysis of functional images with GLM
adwarp @auto_tlrc	Transforming of subjects images to a standard space. Inter-subject register
3dttest	Group analysis of individual statistical maps using T-tests
3dANOVA 1 2 & 3	Group analysis of individual statistical maps using ANOVA
3dRegAna	Group analysis of individual statistical maps using regression
3dFWHM and AlphaSim	Correction for multiple comparisons: Monte Carlo simulation
3dmerge	Correction for multiple comparisons: clustering
3dmaskave	Compute average of all voxels specified by a 3D mask

Table 2.2: Some of the tools provided in AFNI.

In terms of machine learning, a dataset can be represented as a group of observations described by a set of features. For instance, in a MRI set up, given a dataset composed of images extracted from the investigated subjects, the features are the pixels-coordinates of the images and the observations are the subjects encoded as the intensity values related to the feature-pixels. Observations can be grouped into classes that may, or may not be known, for instance "healthy brain" or "pathological brain". Depending on whether these classes are provided to train the machine learning algorithm, or the algorithm is expected to discover new classes, algorithms can be categorized as supervised or unsupervised, respectively.

In biomedical images, machine learning has experienced a spectacular development [de Bruijne, 2016, Madabhushi and Lee, 2016, Rueckert et al., 2016]. The present section will summarize the state of the art of machine learning algorithms used to analyze MR brain images, describing the most widely used feature extraction and dimensionality reduction algorithms, and supervised and unsupervised learning algorithms.

2.4.1 Feature extraction and dimensionality reduction

Although it is possible to use all the available features from a study data set, using too many features can lead to overfitting [Way et al., 2010]. For instance,

this might happen if the algorithm is trained with features that are unrelated, or redundant, to the predicted outcome. Also if the training set does not contain enough observations according to the number of coefficients used by the machine learning algorithm, the trained algorithm will be too specific for the training set, leading to poor performance on new observations [Kuhn and Johnson, 2013, Saeys et al., 2007]. In the last years, this problem has become increasingly important in biomedical imaging, since image resolution has augmented very significantly, leading to more voxels in the images, thus corresponding to more input features [Wang and Summers, 2012].

The first approach to overcome having too many features is called feature selection, which selects a subset of features based on various optimization criteria, reducing in this way the number of considered features [Guyon and Elisseeff, 2003]. Among various feature selection methods, sparse regression models have demonstrated adequate performance in the small sample size problem with different forms of regularization. The Least Absolute Shrinkage and Selection Operator (LASSO) regression model [Tibshirani, 1996] implements an l1-norm to induce sparsity in the regression coefficients for each target response variable independently. To circumvent the limitation of LASSO in the independence in sparsity of the coefficients, Zou and Hastie proposed the elastic net penalty [Zou and Hastie, 2005], which is capable of retaining the sparse property of LASSO and utilizing correlation among predictors simultaneously [Suk et al., 2017].

Peng et al. proposed an additional feature selection method called Minimal Redundancy Maximal Relevance (mRMR) [Peng et al., 2005]. In this method, an optimization is performed affected by two factors: (i) relevance between features and target classes and (ii) redundancy between features. These authors implemented a heuristic framework to minimize redundancy and maximize relevance simultaneously. mRMR showed better performance in many real applications compared with traditional feature selection methods [Wang and Summers, 2012].

Recently, Suk et al., combined two conceptually different methods of sparse regression and deep learning in a unique framework, for the diagnosis and prognosis of Alzheimer's disease or mild cognitive impairment. This method outperformed the competing methods in terms of statistical significance, analyzing an MRI dataset of 805 subjects from the ADNI cohort [Suk et al., 2017].

Dimensionality reduction constitutes another family of approaches used to reduce the number of features employed in a machine learning algorithm. In this approach, the feature space is transformed to extract the useful information, building compact representations from the original data [Wang and Summers, 2012]. Principal Component Analysis (PCA) is one of the most used methodologies of dimensionality reduction that applies orthogonal transformations to the features, in order to create new features that are linearly uncorrelated [Jolliffe, 2002]. Like PCA, Independent Component Analysis (ICA) [Comon, 1994] also looks for a linear transformation,

which can convert the original data to a new linear space. However, ICA looks for transformations that maximize the statistical independence; while in PCA tries to obtain the maximal variance or energy [Wang and Summers, 2012].

In the last decade, new dimensionality reduction techniques using nonlinear transformations have been developed, mainly because numerous data from real applications depict nonlinear structures. Representative nonlinear dimensionality reduction techniques include; (i) local approaches as Locally Linear Embedding (LLE) [Roweis and Saul, 2000] and Laplacian Eigenmaps [Belkin and Niyogi, 2003], and (ii) global approaches such as ISOMAP [Tenenbaum et al., 2000] and Diffusion Map [Coifman et al., 2005a, Coifman et al., 2005b]. Local methods try to preserve the local geometry of the data in low-dimensional space while global approaches tend to give a more faithful representation of the global structure of the data [Wang and Summers, 2012].

2.4.2 Supervised learning algorithms

The goal of supervised learning algorithms is to construct a model that takes an observation x, described by the values of its features, and predict its label y. The model is constructed using the training set, which consists of a dataset of observations x_i where their labels are known and provided to the algorithm [Friedman et al., 2009]. In other words, the algorithm is trained using a labeled training dataset. When applied to classification problems, the inferred outcome for an observation is the related class. For instance, if the algorithm is trained with "healthy brain" and "pathological brain" observations, it will be able to predict the class of new observations of unknown brain state. Supervised classification algorithms can be applied to a wide range of problems in the medical field studied by MRI, such as diagnosis and prognosis [Cuingnet et al., 2011, Sweeney et al., 2014, Basavanhally et al., 2015, Khedher et al., 2015, Armananzas et al., 2016, Patel et al., 2016]. Another application of supervised algorithms is regression, where the outcome of an observation is a continuous function, for instance, the brain performance of a subject on some cognitive task [McLaren et al., 2009, Franke et al., 2010, Rao et al., 2011, Gandjbakhch et al., 2013, Dikaios et al., 2015b, Dikaios et al., 2015a].

An extensive number of supervised algorithms have been developed, for example k-Nearest Neighbors (kNN) [Cover and Hart, 1967], decision trees [Quinlan, 1986], methods based on Linear and Quadratic Discriminant Analysis (LDA/QDA respectively) [Friedman et al., 2009], Naïve Bayes [Domingos and Pazzani, 1997], Support Vector Machines (SVM) [Burges, 1998], Artificial Neural Networks (ANNs) [Friedman et al., 2009], methods based on ensemble learning [Zhang and Ma, 2012], and deep learning [LeCun et al., 2015]. A brief description of this methods and examples of their application to MRI analysis will be addressed below.

k-Nearest Neighbor

k-Nearest Neighbors (kNN) [Cover and Hart, 1967] is a non-parametric method that assigns an observation to the most similar class based on the class of its neighboring observations according to a distance metric. When an unlabeled observation needs to be classified, first its distance to each labeled observation in the training set is computed, using, for example, the Euclidean distance. The observations are ordered in according to the distance and the nearest k neighbors are selected. The predicted class is the most frequent class in these k nearest neighbors.

Note that, in this case, a general model based on labeled observations is not constructed, and there is no training phase. However, all the observations in the training set need to be stored and all the distances need to be recomputed for each new sample to classify. This makes the application of the method computationally intensive and expensive in resources. In addition, it is important to normalize the features in order to have a homogeneous scale, otherwise the features larger scale will have less weight in the similarity computation.

This method has been used in various works to MRI analysis [Vrooman et al., 2007, Besson et al., 2008, El-Dahshan et al., 2010]. For example, Liberman et al. [Liberman et al., 2013] developed a method based on k-NN to classify the enhancing tumor and also the edema-like areas in 59 longitudinal MR data sets of 13 patients with recurrent glioblastoma undergoing bevacizumab therapy.

Decision trees

A decision tree [Quinlan, 1986] is a model encoded as a tree-like graph where each non-leaf node represents a decision test and leafs represent the available classes. In order to classify a new unlabeled observation, the decision test in the root node of the tree is computed, for example if weight feature is greater than 5Kg. Depending on the output, the evaluation process navigates to the children of the node. This process goes on iteratively, until a leaf is reached.

The main advantage of decision trees is that they are easily interpretable and rules in the form of a chain of decisions can be extracted to understand how the samples are classified. Experts can analyze the result to find out the interaction between features that might be previously unknown.

In the MRI analysis field, a variety of reports employing decision trees can be found [Baltzer et al., 2013, Lungu et al., 2016, Bojorquez et al., 2016, Kazerooni et al., 2016, Woitek et al., 2017]. For example, Naik and Patel [Naik and Patel, 2014] used a decision tree to detect and classify tumors in a 124 brain MRI dataset.

Linear and quadratic discriminant analysis

For c classes, the basic idea of Linear Discriminant Analysis (LDA) [Friedman et al., 2009] is to find a linear transformation to a $c-1$ space where samples from different classes are optimally separated (that is, patterns of the same class tend to be close to each other, far away from patterns of other classes). Fisher discriminant analysis [Fisher, 1936] is a particularization of LDA for two-class classification problems (Figure 2.11). Quadratic Discriminant Analysis (QDA) [Friedman et al., 2009] is a more general case that does not assume that the covariance of all the classes are equal using a quadratic transformation of the observation vector instead of a linear one.

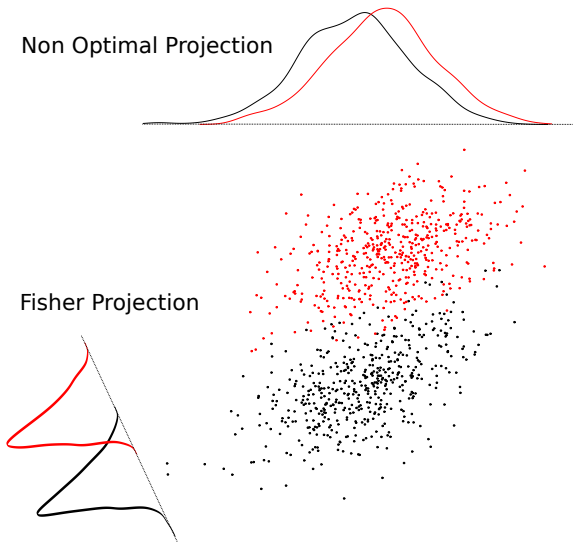

Figure 2.11: Illustration of Fisher discriminant analysis. The method finds the optimal linear projection that maximizes the distances between observations from two classes (dots from the same class are represented with the same color) and minimizes the distances between observations from the same class.

One of the many papers in the literature that uses LDA in MRI analysis is [Fellah et al., 2012]. These authors used a classifier based on LDA to classify between images obtained from pediatric patients with cortical dysplasia, neuroglial tumors and gangliomas. It was shown that classes could be completely separated combining different types of acquired MR images: T_{1w}, T_{2w}, fluid-attenuated inversion recovery, diffusion and magnetic resonance spectroscopy. Additional publications using the LDA approach are [Habes et al., 2013, Xia et al., 2015, Wu et al., 2015b, Mandelkow et al., 2016].

Naïve Bayes

The Naïve Bayes [Domingos and Pazzani, 1997] is a family of classifiers that are trained applying the Bayes theorem with the assumption that the features are independent from each other, which is the naive assumption. This is not usually the case in real examples but the method can be used to get estimates of performance that can be computed efficiently. However, when a small sample is available this method showed a robust behavior [Hand and Yu, 2001].

In the research literature, key-note papers that use a Naïve Bayes classifier to MR images analysis are [Yang et al., 2011, Strumia et al., 2012, Behroozi and Daliri, 2014, Zhou et al., 2015]. Notabely, Griffis et al. [Griffis et al., 2016] trained a Naïve Bayes classifier for automatically identify stroke lesions in T_1-weighted MRI.

Support Vector Machines

Support Vector Machines (SVM) is a classification algorithm based on kernels [Burges, 1998]. In this approach, the samples are mapped to a higher dimensional space in order to archive a better separation of the classes using a hyperplane. Kernel functions are used to transform the samples to the new space because they allow efficient computation of dot products in terms of the original space, which is a required operation in the learning and prediction process. The hyperplane used to separate the classes in the new space is computed by minimization of the classification error and maximization of the margin between the plane and the samples simultaneously.

Kocevar et al. [Kocevar et al., 2016] applies a tuned SVM combined with Radial Basic Function kernel in order to classify multiple sclerosis patients into four clinical profiles from a DTI acquisition. The method was able to accurately classify sixty-four patients with this disease. Additional examples that use SVMs for constructing MRI classification systems are [Tanoori et al., 2011, Graña et al., 2011, El Azami et al., 2013].

Artificial Neural Networks

Artificial Neural Networks (ANNs) [Friedman et al., 2009] is a machine learning method inspired in how neurons process stimulus in the brain. A neural network is composed interconnected entities, called artificial neurons or perceptrons. Each perceptron compute a linear combination of their inputs according to the trained weights followed by a nonlinear function (Figure 2.12 A). To process an observation the feature values are provided to the first layer of the neural network, called input layer, that propagate the computation among different the layers until the output

layer provides the predicted output (Figure 2.12 B).

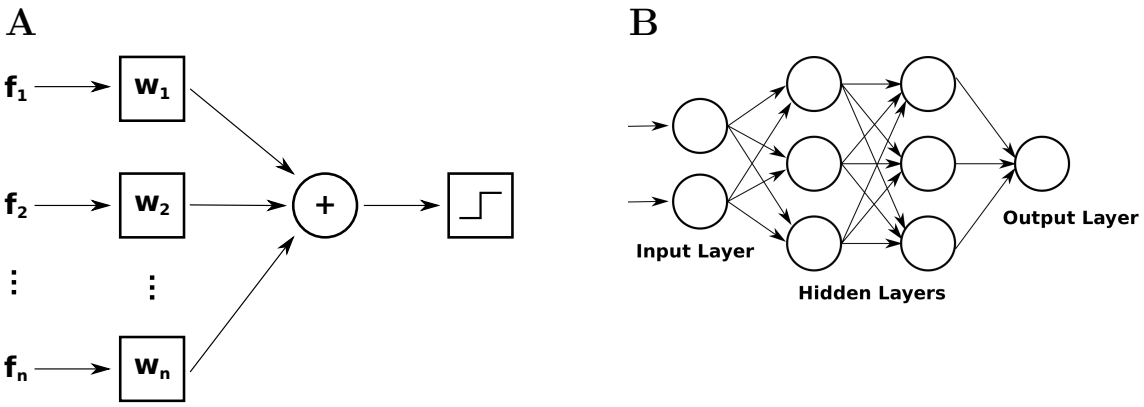

Figure 2.12: Examples of neural network architectures. A: Perceptron. B: Feedforward network with two hidden layers.

There are several types of ANNs, the most widely used is the feedforward network trained with the backpropagation learning algorithm [Weinert, 1979]. El-Sayed et al. [El-Dahshan et al., 2014] used a feedforward back-propagation ANN to classify 101 MRI dataset of human brain into normal and abnormal (with malignant and benign tumors). ANNs have been used to classify MR images in many other reports, such as [El-Dahshan et al., 2010, Dietzel et al., 2012, Yun et al., 2012, Kim and Park, 2017].

Ensemble learning

In ensemble learning [Zhang and Ma, 2012] multiple supervised learning algorithms are trained and combined to resolve a problem. Most learning algorithms are based on optimization problems that can get stuck in a bad local optimum. The idea behind enable learning is that, by combining the outputs of different trained models, the risk of choosing a wrong classifier is minimized.

One approach to enable learning is called bootstrap aggregation or bagging [Breiman, 1996]. In order to prevent the production of many identical trained models, each model is trained with a random resample of the training dataset and the outputs are combined with equal weight. A random forest [Breiman, 2001] is a specific case of bagging, where multiple decision trees are trained with bootstrapping, that proved to have better generalization accuracy than single tree classifiers.

Another approach, used by the AdaBoost algorithm [Freund and Schapire, 1995], builds the classifier set incrementally, where new added classifiers are trained with special focus in the observations that were incorrectly classified in previous iterations of the training process [Schapire, 1990]. This is done by assigning weights to

each observation in the training set, and an error function is computed with the classification error weighted by them. The weights are updated in each iteration, placing more weight in samples that were incorrectly classified. The final model is constructed assigning a weight to each classifier, related to the accuracy in the weighted training set.

An example of AdaBoost algorithm application to MRI analysis is an automated computer-aided diagnosis framework for breast MR images proposed by Lu et al. [Lu et al., 2017], in which AdaBoost is used to classify 2336 breast MRI into malignant tumors, benign tumors and normal tissue. Additional examples that employ ensemble learning approach include; [Simpson et al., 2013, Azmi et al., 2013, Guo et al., 2014, Hirose et al., 2015, Cacha et al., 2016].

Deep learning

Deep learning [LeCun et al., 2015] includes a family of models and training methods capable of learning multiple levels of representation enabling the modelization of complex relations among measured factors. A deep architecture is constructed as a hierarchy of features where higher-level features are defined using lower-level ones. The representation is organized as a set of layers of abstraction where each can be based on different models. One of the main benefits of deep learning techniques is that the feature selection can be handled internally by the abstraction layers [LeCun, 2012, Schmidhuber, 2015]. However, it is important to note that, because the model contains many parameters, deep learning requires huge training datasets and advanced computational technologies, like general-purpose graphical processing units.

The most common model used in deep learning are feed forward neural networks with many hidden layers, called deep neural networks. The classical backpropagation learning algorithm leads to bad results if applied directly to neural networks with many hidden layers [Glorot and Bengio, 2010], especially because the optimization process often is trapped in a local optimum of the optimized function. Many variations of the learning algorithm have been proposed, such as the use of autoencoders to pre-train the network providing a good initialization of the neuron weights [Bengio et al., 2007].

Numerous papers that apply deep learning methods to analyze medical images have been published [Shen et al., 2017]. In this field, one of the most commonly used deep learning algorithm tool is Convolutional Neural Networks (CNNs) [Le Cun et al., 1990, LeCun et al., 1998]. CNNs are feedforward neural networks with a special interconnection layout that is particularly designed to process images. Among recently published works that analyze MRI using CNNs are [Dou et al., 2016, Pereira et al., 2016, Nie et al., 2016, Brosch et al., 2016, Wachinger et al., 2017]. In particular, Zhang et al. [Zhang et al., 2015] proposed a CNNs for isointense infant brain

tissue segmentation, using T_1, T_2 and fractional anisotropy images.

2.4.3 Unsupervised learning algorithms

In unsupervised learning algorithms, the goal is to obtain the natural structure of the dataset [Friedman et al., 2009]. Now the training set is provided to the algorithm without pre-assigned labels and it will group the observations in clusters according to their similarity in a process called clustering. Then, these clusters can be studied to find patterns that are beyond human perception.

Among the most well-known clustering algorithms are k-means [Lloyd, 1982], fuzzy c-means [Bezdek et al., 1984], hierarchical clustering [Friedman et al., 2009], Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [Ester et al., 1996], and clustering based on mixtures of Gaussians trained with Expectation–Maximization (EM) [Dempster et al., 1977, Bilmes et al., 1998]. Clustering methods have been used in medical applications, for example, medical image segmentation and diagnosis [Ortiz et al., 2013, Ahmed et al., 2014, Ryali et al., 2015, Wang et al., 2016, Kaya et al., 2017]. Below, we will review the most important clustering algorithms and their application in the MRI analysis.

k-means and fuzzy c-means

The k-means [Lloyd, 1982] is an iterative algorithm whose objective is to find the optimal partition of the set of observations in k clusters, where each cluster is represented by its means or centroids. First, k centroids are chosen randomly or using some heuristic strategy and each one assigned to a cluster. Next an interactive process of two steps is executed: 1) each observation is assigned to the cluster related to the closest centroid and 2) the centroids are moved to the average of the observations of their cluster. These steps are repeated until the centroids do not change. This is equivalent to minimize the function of the sum of the square of the Euclidean distances of each observation to its centroid.

This algorithm presents several limitations, one is that clusters tends the have spherical or ellipsoidal shapes, that might not fit the data, and makes isolated observations difficult to group. It is also required to guess the correct number of clusters, producing different clustering results every time the algorithm is executed [Garey et al., 1982].

As described above, the k-means algorithm assigns each observation to a concrete cluster, group overlapping is not allowed. The Fuzzy c-means algorithm [Bezdek et al., 1984] is a variant of the k-means algorithm where each observation is related to each group with a membership grade. This algorithm is less likely to fall into

local minimums than k-means.

k-means and fuzzy c-means have been widely used in analysis of MR images [Nguyen et al., 2015, Fontana et al., 2016, Nekooeimehr et al., 2016, Triadyaksa et al., 2016, Gibson et al., 2010, Shen et al., 2011, Chang et al., 2012, Elazab et al., 2015]. For example, Gray et al. [Gray et al., 2011] applied the k-means clustering to identify the contractile and noncontractile tissue compartments on thigh MR images of 34 subjects. On the other hand, Lee et al. [Lee et al., 2016] employed the fuzzy c-means clustering for predicting nodal metastasis in head and neck cancer, analyzing apparent diffusion coefficient maps obtained from diffusion weighed images of 169 lymph node lesions.

Hierarchical clustering

Hierarchical clustering algorithms [Hastie et al., 2009g] build a nested sequence of clusters. Different abstraction levels are defined, where top-level clusters are composed of lower level clusters. These levels can be graphically represented using dendrograms that present a taxonomy of the processed information and allow meaningful navigation in large datasets. According to the procedure used to build the hierarchy, the algorithms are divided in agglomerative and divisive.

Agglomerative algorithms [Murtagh and Contreras, 2012] follow a bottom-up approach. They start considering each observation as an individual cluster. Then, an iterative process is executed where, in every iteration, the most similar clusters are merged. The process stops when all observations are merged into the same top-level cluster.

On the other hand, divisive algorithms follow a top-down strategy, starting with a single cluster that contains all the observations. In every iteration, the cluster is split in two until single observation clusters are created for each observation.

Hierarchal clustering algorithms have been proposed for MRI analysis [Demir et al., 2013, Moreno-Dominguez et al., 2014, Demir and Cetingül, 2015, Marrelec et al., 2015]. For example, Wang and Li [Wang and Li, 2013] proposed a voxel-wise hierarchical clustering framework to analyze whole-brain resting-state fMRI datasets. This framework was evaluated in a resting-state fMRI dataset from 86 adult subjects.

DBSCAN

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [Ester et al., 1996] is a density-based clustering algorithm, where the observations to be

clustered are supposed to follow a concrete probability distribution [Banfield and Raftery, 1993]. In this context, observations are referred as points.

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [Ester et al., 1996] is a density-based clustering algorithm. In this type of algorithm, the groups are the dense subspaces in the space of representation of the observations. This algorithm has two input parameters, *minPoints* and *Eps*. Each observation is defined as a point within that space of representation. The algorithm classifies the points in three types:

- Core point: a point is considered core point, if there are at least *minPoints* within the *Eps* distance.
- Reachable point: a point is reachable from another point if there is a path of core points that connect these.
- Outliers are points that have no other points within the *Eps* distance.

The algorithm visits all the points and create clusters whose points are core, or reachable from a core point, and outliers are not included in any cluster. The intuitive idea is that the clusters are grown connecting density zones of the space, following any kind of shape.

DBSCAN includes many advantages; it is not required to provide the number of target clusters, the clusters can have arbitrary shapes and handles outliers internally. However, for large dimensionality setup the decided value of the parameters *minDistance* and *Eps* can affect the results [Xu and Wunsch, 2005, Kriegel et al., 2011].

This clustering algorithm has been applied in the MRI analysis [Plant et al., 2010, Bandyopadhyay and Paul, 2013, Baselice et al., 2015]. For instance, Roy and Bhattacharyya [Roy and Bhattacharyya, 2015] presented an approach to segment gray and white matters, which use GDBSCAN [Sander et al., 1998], a generalized version of DBSCAN. This approach was tested in a publicly available T_1-weighted coronal human brain MRI dataset.

Gaussians Mixture and Expectation–Maximization

The Gaussians mixture model (GMM) [Banfield and Raftery, 1993] assumes that the observations follow Gaussian mixture statistical distribution. A mixture of k Gaussians is composed of k different Gaussians, each with its μ_i (mean vector), Σ_i a (covariance matrix) and a related probability p_i. The p_i is the mixture weight or prior-probability of membership of an observation to the i-th Gaussian, and must add up to 1.

When this model is used, each observation is supposed to be a random realization of one of the Gaussians in the mixture, where each Gaussian can be considered as a cluster. For instance, in order to simulate new observations, first a random uniform is used to decide which cluster is going to be used according to the prior-probability p_i, then a random realization of the selected Gaussian distribution is computed according to its μ and Σ.

The expectation-maximization algorithm [Dempster et al., 1977, Bilmes et al., 1998] is used to fit the Gaussian mixture parameters where the log-likelihood function is maximized. In other words, the method searches what parameters $\{\mu_i, \Sigma_i, p_i\}_{i=1:k}$ define the distribution that assigns greater probability to the observed dataset. In order to assign an observation to a cluster, it is possible to compute the probability of relationship to the each of the gaussians according to its fitted parameters, observations can then be assigned to the cluster with greatest probability.

Several Gaussian mixture models trained with expectation-maximization have been previously applied in MRI analysis [Lu et al., 2008, Merisaari et al., 2009, Donoso et al., 2010, Oikonomou and Blekas, 2013, Lu et al., 2013]. For example, Simon et al. [Simon et al., 2012] presented a method that automatically delineates and quantifies heterogeneous regions in gliomas in apparent diffusion coefficient maps, using an EM clustering algorithm. This method was evaluated in 10 patients with glioma.

2.4.4 Validation of classification and clustering methods

Classification

Validation strategies applied to a classification problem are aimed to find out if the resulting model will generalize successfully to new data, not used previously to train it. For instance, when the number of parameters of the model is too big in relation to the number of samples used to train the model, the result will just learn by heart the data, and will behave more like a map from the input sample to provided classifications. When a sample that is not within the training set is provided to this model, the predicted class will be randomly assigned, revealing the bad generalization. In this case, the model is said to be overfitted to the data [Erickson et al., 2017].

When a machine leaning algorithm is applied to a problem, knowing if overfitting is happening is of outmost importance. Testing different algorithms and different configurations in a process of refinement, is therefore, a key aspect of machine learning.

In order to be able to validate a model, first, available data is divided into independent training and test sets. Only the training dataset is used to train the model. The quality of the generalization of the trained model can be measured as the prediction accuracy on the test dataset. The way in which observations are partitioned gives rise to different validation strategies.

The most used partition strategies are K-fold cross-validation, leave-one-out, and random sampling. In K-fold cross-validation [Friedman et al., 2009, Arlot et al., 2010], data set is split into k independent subsets. Several different fits and test results are obtained using different combinations of these partitions, each using K-1 of the subsets as training set and the remaining subset as the test set. The accuracy of the validation is computed with the average of the classification accuracies on each test dataset. Leave-one-out [Friedman et al., 2009] is a specific case of K-fold cross-validation where the entire sample but one observation, is used to train each model each time. The main advantage of these approaches is that more data is available to train the model, in the leave-one-out case, all the sample but one observation, which will lead to a better usage of the data. This is a good strategy when the available sample is small. Although many different models result from this process, it should not make a big difference choosing any of them, since they are trained with almost the same sample.

Clustering

In order to validate clustering methods, additional validation metrics have been developed [Halkidi et al., 2001, Halkidi et al., 2002a, Halkidi et al., 2002b, Rendón et al., 2011]. In the literature, these validation metrics are usually categorized into two types: internal and external. Internal metrics are those that evaluate the result of a clustering method based on intrinsic characteristics of the clusters that it obtains. Conversely, the external metrics compare the clusters obtained with the classes previously known to appear in the dataset.

As mentioned above, external evaluation of clustering results are computed using prior knowledge of the input data, usually expert provided classifications. The goal is to test how close are the resulting clusters, to any of the known data structures. These measurements are usually based on methods to validate classification results, instead of measuring the accuracy as the number of points allocated in the correct classes, it is tested if points known to be truly in the same cluster, are allocated together. It is important to note that a flaw in these approaches is that clustering results might find unknown true classes leading to the idea that the result is wrong, when the result is just expressing previously unknown structures. Among the external metrics most referenced in the literature are the following: F-measure [Larsen and Aone, 1999], Nmimeasure [Strehl and Ghosh, 2002], Jaccard index [Halkidi et al., 2002a], Purity [Rendón et al., 2011], and Entropy [Shannon, 1948].

When internal metrics are used to evaluate a clustering result only the data that was clustered is used, in other words no prior knowledge of the structure of the data is used. The idea is to assign a score to a set of clustering results in order to select the best ones. Different measurements have been presented, depending on the score function used. It is important to note here, that this is different from deciding which clustering results are valid, i.e. find true classifications of the data. These procedures are useful when clustering algorithms present free parameters, such as clusters number, and different results are computed testing different configurations. One of the goals is to reduce the overfitting of the resulting classifications. To do so the complexity of the result, usually in terms of number of clusters, is used to penalize the score in several of this measures. Other score criteria, incentive the results with clusters that are well separated. In some cases, it makes sense to compute several different measurements in order to combine their properties. The most mentioned internal metrics are the following: Akaike information criterion (AIC) index [Akaike, 1974], Bayesian information criterion (BIC) index [Raftery, 1986], Calinski-Harabasz index [Caliński and Harabasz, 1974], Davies-Bouldin index (DB) [Davies and Bouldin, 1979], Silhouette index [Rousseeuw, 1987], and Dunn index [Dunn, 1974].

Chapter 3

Experimental data pre-processing

This chapter describes the experimental design, and the image pre-processing steps applied to every dataset acquired. Briefly, we collected a variety MRI datasets to test the methods developed in this thesis. These can be grouped in two categories: datasets from studies of appetite regulation, and datasets from studies of brain tumor characterization.

Among the studies compiled, there are studies in animal models and studies in humans. All animal procedures were approved by the highest institutional ethical committee (INCSIC) and were performed in accordance with Spanish (law 32/2007) and European regulations (EU directive 2010/63/EU, for animal experiments). Human studies were approved by the appropriate Clinical Trial Commission of the Hospital involved. Volunteers signed up an informed consent and confidentiality document, prior to the imaging experiments.

3.1 Appetite studies

3.1.1 Experimental subjects

In the studies designed to evaluate the cerebral response of homeostatic and hedonic components to appetite stimuli, we collected MRI datasets from the brain in rodent models and humans, under feeding or fasting conditions.

Mice or rats, individually identified with ear punches, were housed in cages containing three or four animals per cage under controlled temperature (21-23 ^{o}C) and humidity (47%) conditions with twelve hours light/dark cycles. We used adult male C57BL6/J mice (n=8 group 1 and n=6 group 2, 8-11 weeks, 24±2 g, Charles River, FR) or adult male Wistar rats (n=7, 8 weeks, 220±20 g, Biomedical Research Institute, Comparative Medicine Facilities, ES), both receiving normal chow diet (A04 http://www.safe-diets.com/eng/home/home.html, SAFE Augy, France, 2900 kcal/kg) and water *ad libitum* for one week before they were investigated by MRI, either under the fed, 48h (mice, group 1) and 16h (mice, group 2 and rats) after fasting. All imaging experiments started at 9 a.m. Two animals from different conditions were randomly selected from the cages and examined per

morning, to minimize uncertainties derived from different durations of the circadian rhythms. This procedure was reproduced until all animals were investigated. Design and reporting adhered to the ARRIVE guidelines [Kilkenny et al., 2010].

Human volunteers were investigated in the Magnetic Resonance Unit of the Hospital Nuestra Señora del Rosario (Madrid, Spain, http://www.hospitalrosario.es). Six healthy male participants, aged 24–33, were admitted to the study under the following conditions: healthy clinical trajectory without familiar history of obesity, diabetes or other endocrine disorders; Body Mass Index (BMI) of 18.5–25, corresponding to normal body weight. Volunteers were required to follow a balanced diet (2000–2500 cal/kg) during seven days prior to the basal image acquisitions, with no drinks other than water *ad libitum*, and no medication or abnormal exercise regimes. Specific instructions to follow a balanced diet were provided for each individual at the beginning of the study, the degree of compliance with these being obtained individually before the imaging sessions. All individuals adhered correctly to the protocol. Each volunteer was imaged in two successive conditions; first fed, after one week of balanced diet and second fasted, 24h after food deprivation. Human protocols were approved by the Clinical Trials commission of the Hospital.

3.1.2 MRI sequences

Mice and rats were examined using a 7 T Bruker Avance III scanner equipped with a 90mm gradient coil insert (36 G/cm, max intensity) and mouse head (23 mm), or rat head (38 mm), bird-cage resonators. Anesthesia was induced in the animals inside a Plexiglas® chamber by administering (1mL/min) an isofluorane/oxygen (2% / 99.9%) mixture, and maintained by delivering (1mL/min) isofluorane/oxygen (1% / 99.9%) mixture, through a nose cap during the MRI protocols.

T_2-weighted (T_{2w}) spin-echo anatomical axial images were acquired using the rapid acquisition with relaxation enhancement (RARE) sequence with the following parameters: repetition time (TR) = 3200 ms, echo time (TE) = 60 ms, RARE factor = 8, averages (Av) = 3, field of view (FOV) = 21 mm x 21 mm (mice) / 32 mm x 32 mm (rats), acquisition matrix (Mtx) = 256 x 256 corresponding to an in-plane resolution of 0.08 mm x 0.08 mm (mice) / 0.12 mm x 0.12 mm (rats), number slices = 3, and slice thickness = 1.5 mm.

Diffusion Weighted Imaging (DWI) was performed using the Stejskal-Tanner spin-echo sequence in three orthogonal directions (left-right L-R, anterior-posterior A-P, and head-foot H-F) as defined by the orientation of the diffusion encoding gradient, using a 4 shot EPI read-out train. Acquisition conditions were: diffusion gradient duration (δ) = $4ms$, diffusion gradient separation (Δ) = $20ms$, TR = 3000 ms, TE = 51 ms, Av = 3, FOV = 38 mm x 38 mm (mice) / 32 mm x 32 mm (rats), Mtx = 128 x 128, corresponding to an in-plane resolution of 0.30

mm x 0.30 mm (mice) / 0.25 mm x 0.25 mm (rats), number of slices = 3 and slice thickness = 1.5 mm. For the eight mice of group 1, we obtained five low b-values ($10 \leq b \leq 100 s/mm^2$) with two basal images (without diffusion gradient application), and six high b-values ($200 \leq b \leq 1800 s/mm^2$) acquisitions with two basal images. In addition, we obtained six b-value acquisitions with two basal images for the six mice of group 2, and three b-value acquisitions with one basal image for the seven rats, both in the range $200 \leq b \leq 1200 s/mm^2$.

Human volunteers were investigated using a GE Medical Systems 1.5 T horizontal-bore superconducting magnet, equipped with a ^1H quadrature head resonator. Image acquisitions were medically supervised by the neuroradiology staff of the hospital. Briefly, a plane containing the hypothalamus was selected using T_{2w} scout images and multi b-value diffusion weighted images were acquired using seven b-values ($200 \leq b \leq 1200 s/mm^2$). T_{2w} spin echo anatomical images were acquired then in every subject, using a rapid acquisition sequence with Fast Relaxation Fast Spin Echo (FRFSE) in coronal orientations. Acquisition parameters were: TR = 3000 ms, TE = 87.08 ms, Echo number = 1, Echo Train Length = 8, Av = 4, FOV = 240 mm x 240 mm, Mtx = 512 x 512, corresponding to an in-plane resolution of 0.47 mm x 0.47 mm, number of slices = 15 and slice thickness = 3 mm. DWI was performed using the following acquisition conditions: TR = 5000 ms, TE = 92.8 ms, Av = 4, FOV = 240 mm x 240 mm, Mtx = 256 x 256, corresponding to an in-plane resolution of 0.94 mm x 0.94 mm, number of slices = 15 and slice thickness = 3 mm.

3.1.3 Pre-processing

We applied three consecutive image pre-processing steps: 1) Slice selection and brain extraction, 2) Bias field correction, and 3) Intensity normalization. These pre-processing steps were implemented using in-house written scripts under the MATLAB (The Mathworks, Natick, US) environment.

Slice selection and brain extraction

In order to analyze simultaneously the activation of homeostatic and hedonic centers, we selected, from the available MRI slices of the brain in the three species, the slice containing the hypothalamus and coplanar hedonic structures. This was conveniently identified using anatomical coordinates as the slice containing the hypothalamus (located above the hypophysis), the third (3v) and dorsal third ventricles (D3v) in rodents or the lateral and third ventricles in humans (Figure 3.1). The selected slice corresponded to coordinates -1.46 mm from Bregma in mice and -3.01 mm from Bregma in rats [Paxinos and Franklin, 2001, Paxinos and Watson, 2004]. In humans, the selected slice contained the right (-4 mm, 2 mm, -6 mm)

and left (-4 mm, -6 mm, 3 mm) hypothalamus and coplanar structures, in Talairach coordinates [Talairach and Tournoux, 1988].

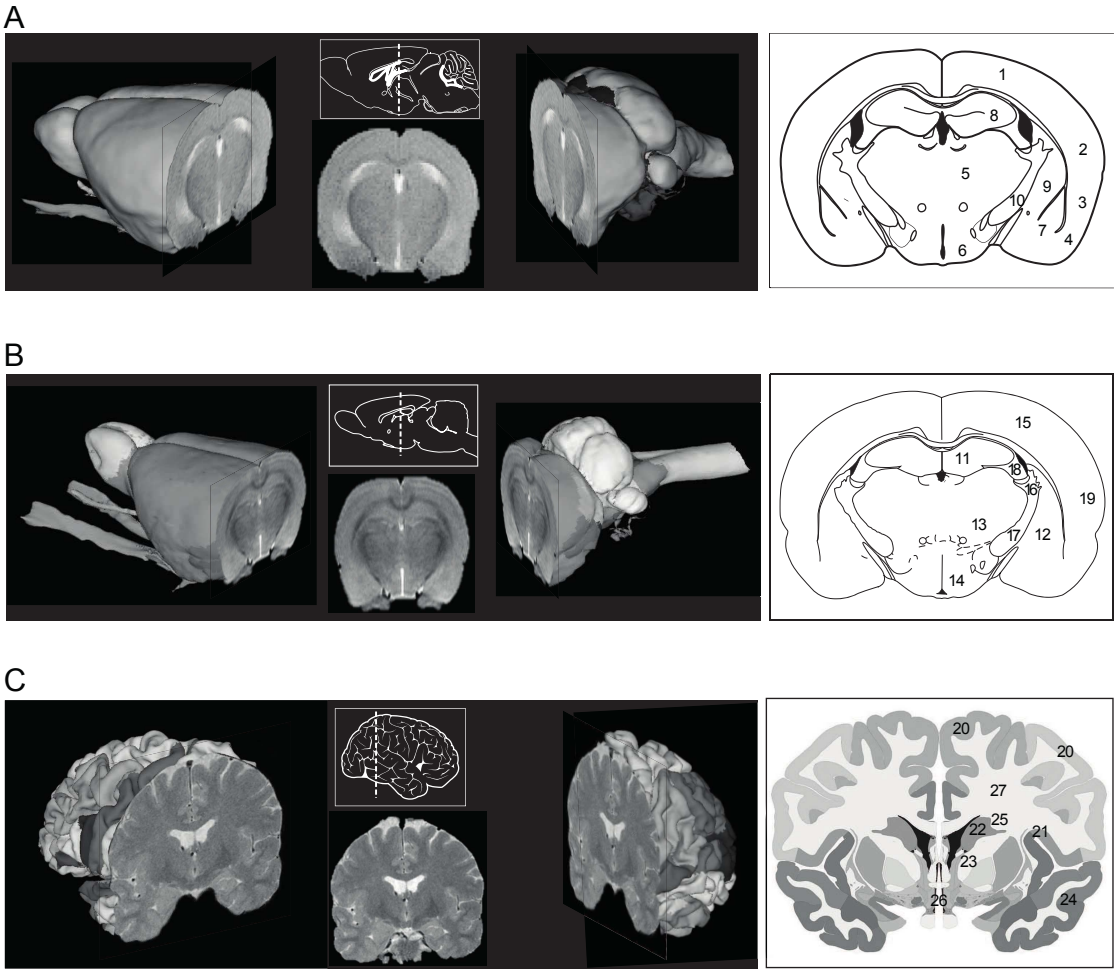

Figure 3.1: Anatomical location of the slice selected in the rodent and human brain. A: Mice. B: Rats. C: Humans. Main cerebral structures contained in the selected slice appear indicated by numbers. 1: somatosensory cortex, 2: auditory and temporal association areas, 3: ectorhinal and peririhnal areas, 4: pyriform layers, 5: thalamus, 6: hypothalamus, 7: amygdala, 8: hippocampus, 9: caudoputamen, 10: globus pallidus. 11: hypoccampus, 12: striatum, 13: thalamus, 14: hypothalamus, 15: neocortex, 16-17: stria terminalis and corticofugal pathways, 18: fimbria of the hypoccampus, 19: basal forebrain region. 20: precentral and postcentral gyrus, 21: parietal operculum and short insular gyri, 22: globus pallidus, 23: amygdaloid complex, 24: superior, median and inferior temporal gyrus, 25: putamen, 26: hypothalamus, 27: white matter of the fore brain.

Each pixel from the selected images can be defined as an intensity vector containing components for each b-value and each orthogonal direction. Then, for mice of group 1, intensity vectors contained thirty intensity values per pixel, corresponding to eleven b-values multiplied by three orthogonal directions acquired. For mice of group 2, intensity vectors contained eighteen intensity values per pixel, corresponding to six b-values multiplied by three orthogonal directions acquired. For humans,

intensity vectors contained twenty-one intensity values per pixel, corresponding to seven b-values multiplied by three orthogonal directions acquired. For rats, intensity vectors contained nine intensity values per pixel, corresponding to the three b-values acquired in three orthogonal directions. Therefore, each subject is defined by two vectors of pixel intensities, for the fed and fasted conditions, respectively. In the following, we shall refer to each one of these vectors, as subject-condition vector.

In order to isolate the brain tissue from the contributions of surrounding anatomical structures (muscle and skull), we used the Pulse-Coupled Neural Networks technique [Chou et al., 2011] for rodents and the Brain Extraction Tool (BET) [Smith, 2002] for humans.

Bias field correction

A polynomial regression of second order was used to correct the non-uniform intensities in images that may have been caused by imperfections in the magnetic field.

Intensity normalization

Different intensity normalization methods were used: a) Intensity normalization based on the basal image, and b) Intensity normalization based on the thalamus region.

1. Intensity normalization based on the basal image. The pixel intensity vectors from each subject-condition were normalized as

$$S'_i(b,d) = S_i(b,d)/S_i(0,0) \quad 1 \leq i \leq n \quad 1 \leq d \leq 3 \tag{3.1}$$

   where $S'_i(b,d)$ refers to the normalized intensity value of i^{th} pixel, acquired with b-value b in d^{th} direction, $S_i(b,d)$ is the intensity value of i^{th} pixel prior to normalization, acquired with the same b-value and direction, $S_i(0,0)$ is the intensity of the i^{th} pixel acquired with $b = 0$, and n is the number of pixels considered in the image. Note that relative intensity of the basal image for $b = 0$, $S_i(0,0)$, is independent of the diffusion gradient direction.

2. Intensity normalization based on the thalamus region. Thalamic region from the basal image at $b = 0$ was used as reference. For each subject-condition, we calculated the average (μ) and standard deviation (σ) of the pixel intensity values from the basal image of the thalamus, and then the normalized subject-condition pixel intensities $S'_i(b,d)$ as

$$S'_i(b,d) = (S_i(b,d) - \mu)/\sigma \quad 1 \leq i \leq n \quad 1 \leq d \leq 3 \tag{3.2}$$

Finally we applied Principal Component Analysis (PCA) [Friedman et al., 2009] to reduce dimensionality and increase Signal-to-Noise Ratio (SNR). The principal components explaining more than 99% of the total variance in data were used in the subsequent analysis.

3.2 Tumor studies

We also tested the developed methods using MRI datasets of brain tumor in rodent models and human patients.

3.2.1 Experimental subjects

Low and high-grade glioma rat model

Albino Wistar rats (Rattus norvegicus) were used (body weight (b.w.) 230 ± 20 g). The animals were kept in cages in a controlled room, with a 12-h cycle of light and darkness, temperature of 22°C ± 2°C and water and food access *ad libitum*. A group of twenty tumor-bearing rats were employed (n = 20), divided in two subgroups of low-grade (n = 10) and high-grade (n = 10).

Authenticated rat glioma C6 cells were obtained from the American Type Culture Collection (ATCC number CCL-107) (Manassas, VA, USA). Cells were grown in Dulbecco's modified Eagle´s medium containing 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid to which, 10% fetal bovine serum (Gibco®, Thermo Fisher Scientific, Inc., Waltham, MA, USA) and antibiotics, were added. The animals were intraperitoneally (i.p.) anaesthetized with a mixture of ketamine hydrochloride (75 mg/kg b.w.) and medetomidine hydrochloride (0.5 mg/kg), and placed in a stereotaxic device, where 10^5 C6 cells in 10 μL of medium were injected in the right caudate nucleus [Benda et al., 1971]. After surgery, the rats were induced to recover with atipamezole hydrochloride (5 mg/kg b.w.) administrated subcutaneously (s.c.). They also received meloxicam s.c. as analgesia (0.5 mg/kg b.w.), during the following 3 days [Pérez-Carro et al., 2014].

Low and high-grade glioma in human patients

We chose a set of subjects from The Cancer Genome Atlas Low Grade Glioma (TCGA-LGG) and The Cancer Genome Atlas Glioblastoma Multiforme (TCGA-GBM) collections, stored on The Cancer Imaging Archive (TCIA) [Clark et al., 2013].

As explained in the description of these collections (http://www.cancerimagingarchive.net, May 2017), most of the images were acquired as part of routine care in many sites all over the world, and not as part of a controlled research study or clinical trial. Therefore, the image datasets were very heterogeneous with respect of scanner modalities, manufacturers and acquisition protocols. We selected 17 patients from collection TCIA-LGG and 18 patients from TCIA-GBM collection (table 3.1), taking into account that they agreed in scanner characteristics, manufacturers and acquisition protocols.

Gender differences in a high-grade glioma rat model

Albino adult Wistar rats (Rattus norvegicus) were used (body weight 230 ± 20 g). The animals were kept in cages in a controlled room, with a 12-h cycle of light and darkness, temperature of 22°C ± 2°C and water and food access *ad libitum*. A group of twelve tumor-bearing rats were employed (n = 12), divided in two subgroups of male (n = 6) and female (n = 6) individuals [Pérez-Carro et al., 2014]. Tumors were implanted following the same procedure described above for the low and high-grade glioma rat model.

Anti-angiogenic therapy response in high-grade glioma

Ten CD-1 mice were treated with B20-4.1.1- Mouse Mab anti-VEGF (Genentech SF, CA) 5mg/Kg IP b.i.w for 4 weeks, starting when the tumor grew over 1mm in size. Among a total of ten treated mice, 6 responded to treatment and 4 did not respond, according to RECIST guidelines [Therasse et al., 2000].

Tumor implantation was done following the same procedure explained in low and high-grade glioma rat model, but in this case, we used authenticated GL261 cells ($10\mu l$ culture media) from the American Type Culture Collection (ATCC number CRL-1887) (Manassas, VA, USA).

3.2.2 MRI sequences

Low and high-grade glioma rat model

In vivo MR experiments in rodents were performed on a Bruker Pharmascan system (Bruker Medical GmbH®, Ettlingen, Germany) using a 7.0-T horizontal superconducting magnet, equipped with a 38-mm diameter ^1H selective birdcage resonator and a gradient insert with 90-mm diameter and a maximum intensity of 360 mT/m [Pérez-Carro et al., 2014]. All data were acquired using the ParaVision

Collection	Patients ID	MRI acquisition date
TCIA-LGG	TCGA-DU-5849	05/04/1995
	TCGA-DU-5852	09/07/1995
	TCGA-DU-5871	06/12/1994
	TCGA-DU-5874	10/05/1995
	TCGA-DU-6407	09/01/1994
	TCGA-DU-6542	08/05/1996
	TCGA-DU-7018	20/12/1991
	TCGA-DU-7019	08/09/1994
	TCGA-DU-7302	03/12/1991
	TCGA-DU-7304	25/03/1993
	TCGA-DU-7306	12/05/1993
	TCGA-DU-8162	29/10/1996
	TCGA-DU-8163	19/11/1996
	TCGA-DU-8164	11/01/1997
	TCGA-DU-8165	05/02/1997
	TCGA-DU-8167	02/04/1997
	TCGA-DU-8168	03/05/1997
TCIA-GBM	TCGA-06-0121	25/03/2003
	TCGA-06-0122	14/09/2004
	TCGA-06-0125	05/11/2001
	TCGA-06-0130	18/06/2002
	TCGA-06-0133	28/07/2005
	TCGA-06-0137	23/09/2003
	TCGA-06-0138	25/11/2002
	TCGA-06-0139	10/06/2004
	TCGA-06-0139	09/11/2004
	TCGA-06-0143	11/02/2005
	TCGA-06-0143	03/11/2005
	TCGA-06-0185	16/11/2007
	TCGA-06-0192	08/04/2005
	TCGA-06-0241	30/08/2005
	TCGA-06-0648	20/01/2006
	TCGA-06-0878	14/10/2006
	TCGA-06-1801	27/09/2007
	TCGA-06-5413	17/06/2008

Table 3.1: Selected patients from the TCIA-LGG and TCIA-GBM collections. Selected considering similarity in scanner characteristics, manufacturers and acquisition protocols. The second and third columns contain identifiers of the selected patients and date of MRI acquisition, respectively.

5.1 software (Bruker Medical Gmbh®) operating on a Linux platform. Anesthesia was initiated in an induction box through inhalation of oxygen (1 L/min) containing 3% to 4% of isofluorane and maintained during the experiment by employing a nose mask with a flow of 1% to 1.5% of isofluorane in O_2. The animals were placed in a heated probe, which maintained the body temperature at approximately 37°C. The physiological status of the rats was monitored by a gating system designed for small animals (SA Instruments, Inc., Stony Brook, NY, USA), using the respiratory rate and body temperature.

T_{2w} spin-echo images were obtained with a RARE sequence and the following parameters: TR = 3000 ms, TE = 60 ms, Av = 3 and RARE factor = 8, Mtx = 256 × 256, FOV = 35 mm × 35 mm, corresponding to an in-plane resolution of 0.14 mm × 0.14 mm, slice thickness = 1.5 mm and five slices in axial orientation.

Magnetization transfer (MT) images [Wolff and Balaban, 1989] were acquired using TR = 2500 ms, TE = 10 ms, Av = 1, FOV = 35 mm x 35 mm, Mtx = 128 × 128, corresponding to an in-plane resolution of 0.27 mm × 0.27 mm, slice thickness = 1.5 mm and five slices in axial orientation. The imaging set acquired with an MT pulse turn on employs a radiofrequency pulse train (N = 50) of bandwidth = 550 Hz, length = 5 ms, power = 5.5 μT and offset = 1500 Hz.

Low and high-grade glioma in human patients

The pre-gadolinium T_{1w} and post-gadolinium T_{1w} Fluid Attenuation Inversion Recovery (T_{1w} FLAIR) images were acquired with scanners from the same manufacturer (GE Medical Systems) and mainly using the SIGNA EXCITE v12 software or the GENESIS_SIGNA v8 software. Pre-gadolinium T1w images were obtained with following parameters: 2497.60 <TR<3285.60 ms, 6.36 <TE< 8.10 ms, Echo number = 1, Echo Train Length = 4 and Av = 1, with the geometry parameters: Mtx = 512 × 512, FOV = 240 mm × 240 mm corresponding to an in-plane resolution of 0.47 mm × 0.47 mm, 2.5 <slice thickness< 5 mm, and 32-84 slices in axial orientation. In addition, post-gadolinium T_{1w} FLAIR images were obtained with the following parameters: 2039.90<TR<3379.60 ms, 6.36<TE<8.10 ms, Echo number = 1, Echo Train Length = 4 and Av = 1, with the geometry parameters: Mtx = 512 × 512, FOV = 240 mm × 240 mm corresponding to an in-plane resolution of 0.47 mm × 0.47 mm, 2.5<slice thickness<5 mm and 32-42 slices in axial orientation.

Gender differences in a high-grade glioma rat model

In vivo MR experiments in rats were carried out on the same scanner and under the same conditions explained above in the Low and high-grade glioma models. Likewise, T_{2w} images were obtained following the same acquisition protocol indicated

above for low and high-grade glioma rat models.

Diffusion tensor imaging (DTI) [Basser et al., 1994] studies were performed by running a Stejskal-Tanner sequence with a single-shot echo-planar readout gradient, where TR = 3000 ms, TE = 40 ms, Av = 4, FOV = 35 mm x 35 mm, Mtx = 128 × 128 with an in-plane resolution of 0.27 mm × 0.27 mm, Δ = 16 ms, δ = 4 ms, one basal image (without diffusion gradient application) and two b factors of 300 and 1,400 s/mm^2, corresponding to diffusion gradient strengths of 33% and 71%, respectively, as applied in seven directions.

Anti-angiogenic therapy response in high-grade glioma

Anti-angiogenic treatment response to bevacizumab [Norden et al., 2009, Khasraw et al., 2014] was evaluated in ten mice bearing implanted GL261 tumors on a 7T MR scanner. MRI studies included DWI and post-gadolinium T_{1w} images performed sequentially at 3/4 day intervals until completion of treatment or obvious, treatment independent, tumor growth.

DWI was performed using the following acquisition conditions: EPI sequence, 3 directions, b-values = 100, 400 and 1000 s/mm^2, TR = 3000 ms, TE = 51 ms, Av = 4, Mtx = 128 × 128, FOV = 35 mm × 35 mm resulting in an in-plane resolution of 0.27 mm x 0.27 mm, slice thickness = 1 mm and six slices in axial orientation.

Contrast-enhanced $T_{\text{1-weighted}}$ (CE-T1w) SE images after i.v. administration of 0.3 M Gd-diethylene triamine pentaacetic acid (DTPA) (from Magnevist®, Bayer, Whippany NJ, USA), at a dose of 0.3 mmol/kg b.w., were acquired using TR = 337 ms, TE = 11 ms and Av = 2. T_{1w} images were acquired with the geometry parameters: Mtx = 256 × 192, FOV = 22 mm × 22 mm resulting in an in-plane resolution of 0.11 mm x 0.09 mm, slice thickness = 1 mm and 15 slices in axial orientation.

Additionally, T_{2w} spin-echo anatomical axial images were acquired with the RARE sequence using the following parameters: TR = 3000 ms, TE = 60 ms, RARE factor = 8, Av = 3, FOV = 22 mm x 22 mm, Mtx = 256 x 192 corresponding to an in-plane resolution of 0.11 mm x 0.09 mm, number slices = 15, and slice thickness = 1 mm.

3.2.3 Pre-processing

Two consecutive image pre-processing steps were applied: 1) Slice selection and tumor delimitation, and 2) Intensity normalization. In the same way as in the case of appetite studies, these pre-processing steps were implemented using in-house

written scripts under the MATLAB (The Mathworks, Natick, US) environment.

Slice selection and tumor delineation

We selected, from the available MRI slices of the brain in each dataset of tumor studies, the slice containing the largest tumor area (Figure 3.2).

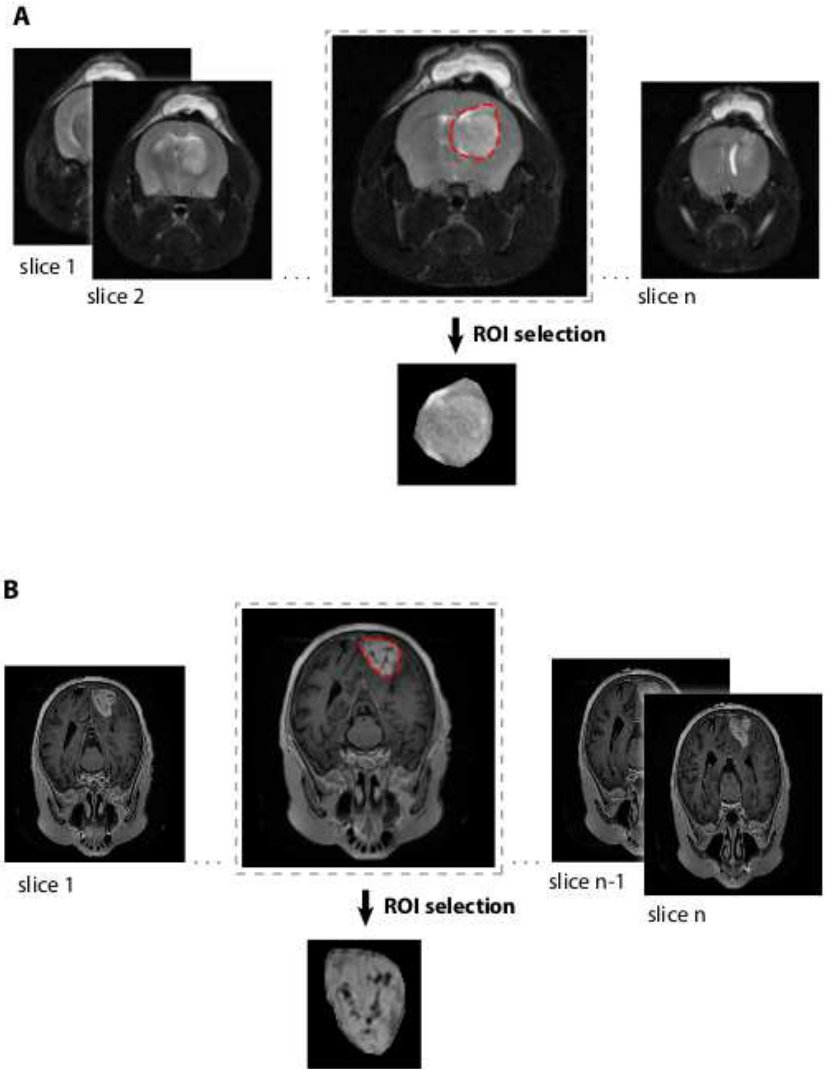

Figure 3.2: First pre-processing step for tumor studies datasets. The slice containing the largest tumor area is selected, and the tumor area is delimitated and extracted for subsequent analysis.

Similar to the appetite studies, in tumor studies each pixel from the selected images can be defined as an intensity vector containing components for each image acquired in the study. Then, for low and high-grade glioma rat model, intensity vectors contain two intensity values per pixel, one is the intensity value in MT

image and the other is the intensity value in T_{2w} image. For low and high-grade glioma in human patients, intensity vectors contain two intensity values per pixel also. One is the intensity value in pre-gadolinium T1w image, and the other, is the intensity value in post-gadolinium T_{1w} FLAIR image. For gender differences in a high-grade glioma rat model, intensity vectors contained fourteen intensity values per pixel, corresponding to two b-values multiplied by seven directions acquired. For Anti-angiogenic therapy response in high-grade glioma, intensity vectors contained nine intensity values per pixel, corresponding to three b-values multiplied by three directions acquired. Therefore, each subject is defined by two vectors of pixel intensities, one for each condition (low-grade and high-grade, female and male, positive response and negative response). We shall refer to each one of these vectors as subject-condition vector.

Intensity normalization

To normalize intensities, we selected for each dataset and subject-condition, a reference ROI from the healthy brain, depicting the most homogeneously possible intensity values. In DWI, for each subject-condition, we calculated the average (μ) and standard deviation (σ) of the pixel intensity values from the basal image of selected ROI, and then the normalized subject-condition pixel intensities, as indicated by equation 3.2. For other image types, we calculated μ and σ in the reference ROI from the selected slice.

As was done in the appetite studies (Section 3.1.3), Principal Component Analysis was used to reduce dimensionality and increase SNR.

Chapter 4

Supervised learning

This chapter implements new supervised learning algorithms for the analysis of MRI datasets as obtained from a small number of subjects. The proposed methods evolved from the idea of considering each pixel as a pattern, rather than from the conventional notion of considering each image as a pattern. We illustrate this approach, presenting first the Fisher Maps methodology, allowing for the non-invasive and comprehensive visualization of the cerebral appetite circuitry, through the automatic analysis of DWI datasets. Then, we expand this methodology to the classification of complete images, using the set of predictions obtained for all its pixels. We evaluate and validate the new classification strategies using two different types of studies: cerebral appetite stimulation and glioma grading.

4.1 Introduction

Appetite is a vital cerebral response integrating a large variety of endocrine and intracerebral signals determining ultimately the global energy balance between food intake and energy expenditure [Morton et al., 2006, Williams and Elmquist, 2012]. Specific mechanisms underlying the cerebral appetite response remain, however, insufficiently understood, particularly *in vivo*. Notably, disturbances in these mechanisms are thought to underlie global and daunting pandemics as obesity or diabetes [Schwartz and Porte, 2005, Berthoud and Morrison, 2008, Moheet et al., 2015]. Neuroimaging methods have contributed substantially to improve our knowledge on cerebral appetite regulation. In particular, Diffusion Weighted Imaging (DWI) has been proposed as a useful methodology to detect areas of cerebral activation and neuronal connectivity based on the alterations on water diffusion parameters and their anisotropy [Le Bihan, 2003]. Interestingly, DWI revealed small but significant alterations in water diffusion parameters of the hypothalamus and its different nuclei upon fasting, but the contribution of additional cerebral structures, as the cortico-limbic system, was not addressed [Lizarbe et al., 2013, Lizarbe et al., 2015]. The underlying neurophysiological DWI interpretations remained, however, uncertain [Le Bihan, 2012, Norris, 2001, Syková and Nicholson, 2008, Ackerman and Neil, 2010], relying frequently in simplified models of water diffusion through the brain tissue and the *a priori* selection of specific cerebral regions previously reported to be involved in appetite regulation [Yablonskiy and Sukstanskii, 2010].

We found it, therefore, of interest to investigate here a model-independent approach to analyse automatically DWI datasets, improving the restrictions imposed by the phenomenological models of diffusion and the *a priori* selection of specific regions of interest for analysis. The first part of this Chapter reports on the implementation of this methodology to explore the effects of appetite stimulation in the brain of rodents and humans, comparing the results with those derived from the classical mono-exponential model fittings of the same datasets.

We then extend the methodology to the classification of complete images, illustrating its performance with the non-invasive classification of appetite stimulation, glioma grade and sexual dimorphism in rodents and humans, using different types of MR images. Gliomas are primary brain tumors of the central nervous system and the most frequent in adults [Raizer et al., 2015]. Despite enormous progress in basic cancer research, including the advances in their molecular or cellular determinants and their diagnosis and treatment, gliomas remain among the highest mortality brain tumors [Stylli et al., 2015, Puttick et al., 2015]. In recent years, non-invasive imaging techniques, including MRI, have played a crucial role in glioma diagnosis and prognosis, supporting elegantly the development of new therapeutic strategies [Barajas Jr and Cha, 2014, Dhermain, 2014]. Many machine learning algorithms have been implemented to analyse automatically these images, improving considerably diagnostic or therapeutic decisions [Zöllner et al., 2012, Svolos et al., 2013, Ma and Song, 2013, Schäfer et al., 2013, Jung et al., 2014, Wu et al., 2015a, Ranjith et al., 2015]. Despite these valuable efforts, the non-invasive identification of glioma grade, remains today as a fundamental medical challenge to improve rapid and adequate handling of glioma patients. To overcome this limitation, we propose here a classification method to identify automatically low and high-grade gliomas in a rat model and in human patients, from MRI images. Additionally, we extended this methodology to investigate the sexual dimorphism in a high-grade glioma rat model, classifying automatically the pathology between male and female carriers [McKinley et al., 2000, Sun et al., 2012, Pérez-Carro et al., 2014].

4.2 Fisher maps

4.2.1 Methods

Figure 4.1 illustrates the experimental approach implemented for the appetite studies, using the mouse brain model as an example. Briefly, we acquired DWI datasets from six mice (Figure 4.1 A) under the fed (blue) or fasted (red) conditions applying two independent image processing strategies for the DWI dataset analysis, either the classical mono-exponential model fitting (Figure 4.1 B) yielding ADC maps (Figure 4.1 C), or the proposed Fisher Maps method (Figure 4.1 D) generating Appetite Index maps (Figure 4.1 E).

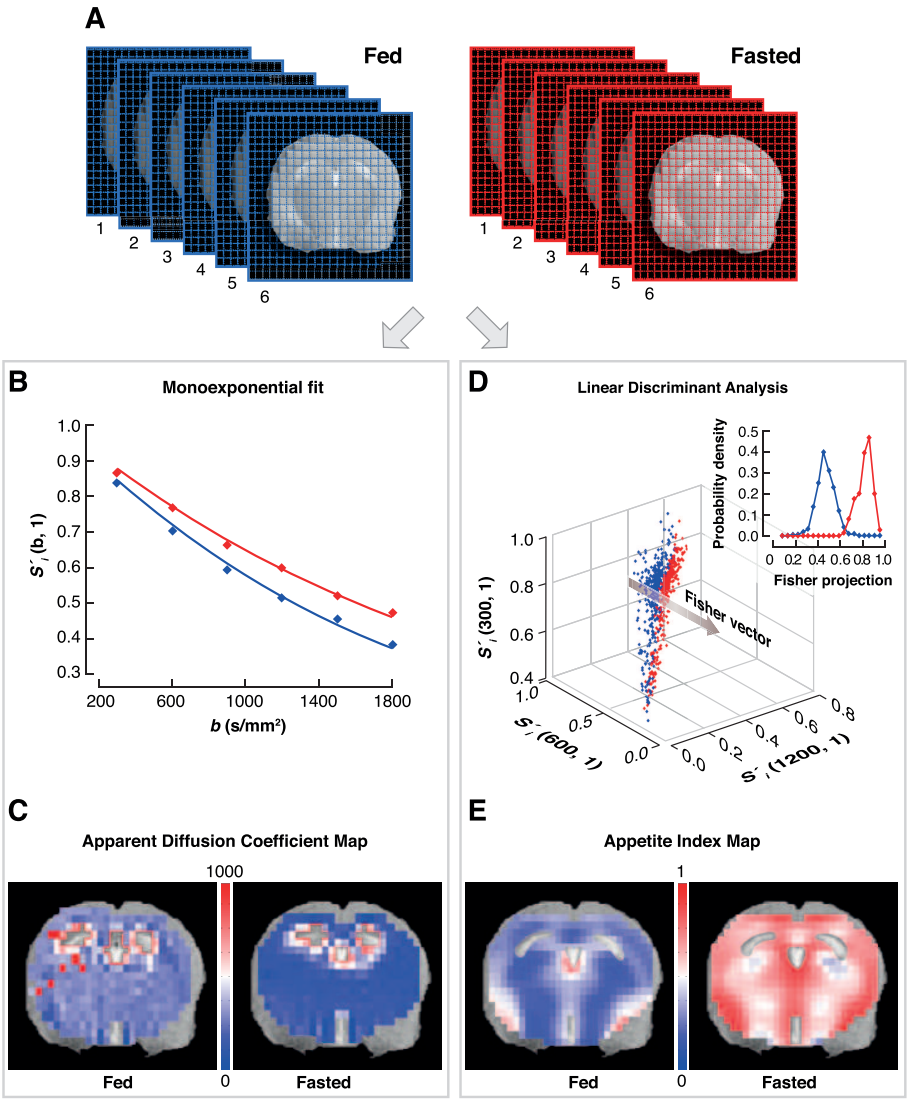

Figure 4.1: Overview of the methodology. A: DWI datasets acquired from the brain six mice under the fed (blue) or fasted (red) conditions. B: Representative mono-exponential fits from the same pixel of the same mouse under the fed or fasted conditions. C: ADC maps from the mouse brain under the fed (left) or fasted (right) conditions. The color bar indicates the full range of ADC values. D: Linear Discriminant Analysis determination of the Fisher vector (grey arrow) providing optimal discrimination between the fed (blue) and fasted (red) pixel clouds. The projection of the pixel intensity over the Fisher vector can be treated as an appetite score, termed Appetite Index (AI). Inset D: Histograms of the projections of fed and fasted pixels over the Fisher vector. E: Appetite Index maps generated from mouse brain under the fed or fasted conditions. The color bar indicates the full range of Appetite Index values. The ventricular spaces were not included in the analysis (white spaces in C and E).

In the model-dependent approach (Figure 4.1 B), the classical mono-exponential model was fitted to the six b-value intensities in three orthogonal directions for each pixel, obtaining the average ADC value for every pixel and direction, for every subject, and feeding condition. Figure 4.1 C presents an illustrative ADC map

(Left-Right direction) of a representative mouse brain, under the fed (left) or fasted (right) condition.

The Fisher Maps method (Figure 4.1 D) used Linear Discriminant Analysis (LDA) [Friedman et al., 2009] to find the Fisher vector w that optimally separates the fed (blue) and fasted (red) pixel clouds in a multidimensional space of six b-values and three orthogonal directions. This approach is graphically demonstrated with the data of a representative mouse brain in a three dimensional space (b-values of 300, 600 and 1200 s/mm^2 in Left-Right direction), illustrating how the fed and fasted pixels can be separated by the Fisher vector direction w (grey arrow). Figure 4.1 D (inset) illustrates the histograms of the projections of individual pixels for both classes, showing an adequate separation between the fed and fasted pixel clouds. From these histograms, the probability that a pixel with a given projection corresponds to an image obtained in the fasted condition, can be calculated straightforwardly. We termed this probability the "Appetite Index" (AI) of the pixel. Figure 4.1 E depicts the resulting AI values for every pixel, as located in the original coordinates, rendering the AI map for each fed (left) or fasted (right) pixel in the investigated mouse brain section.

The objective of this approach is to discriminate the cerebral regions whose activations are related to the fasting stimulus. To this end, we applied the LDA algorithm to find the Fisher vector w that optimally separates the fed from fasted pixels. Each pixel was then projected over the Fisher vector using the expression

$$z_i = \sum_{b,d} w_{b,d} \cdot S_i'(b,d) \quad 1 \leq i \leq n \quad 1 \leq d \leq 3 \tag{4.1}$$

where z_i is the projection of i-th pixel. The vector z is the Fisher projection, which contain the projected intensities of all pixels.

Finally we calculated the Appetite Index of the pixel i, AI_i, as

$$AI_i = P(\text{fasted}|z_i) \tag{4.2}$$

where $P(\text{fasted}|z_i)$ indicates the probability that a pixel with projection z_i belongs to a fasted image. Therefore, AI may be interpreted as an appetite score.

Using Bayes' Theorem and the fact that $P(\text{fasted}) = P(\text{fed}) = 1/2$ in our datasets, we get

$$AI_i = \frac{P(z_i|\text{fasted})P(\text{fasted})}{P(z_i|\text{fasted})P(\text{fasted}) + P(z_i|\text{fed})P(\text{fed})} \tag{4.3}$$

$$= \frac{P(z_i|\text{fasted})}{P(z_i|\text{fasted}) + P(z_i|\text{fed})} \tag{4.4}$$

where he probability densities $P(z_i|\text{fasted})$ and $P(z_i|\text{fed})$ are calculated from the histograms of the pixel projections for each condition.

To validate the generalization of this approach, we implemented a leave-one-subject-out strategy [Friedman et al., 2009], consisting in a series of cross-validation rounds, where a different Fisher projection vector is obtained with the images of all subjects except one, both in fed and fasted images for that subject, then used as "test subject". In each round, the calculated Fisher projection vector is applied to the images of the corresponding test subject, obtaining the AI score for each pixel under the fed and fasted conditions. Once all leave-one-out subject rounds were completed, fed and fasted pixel populations were created accumulating the AI pixel values from the fed and fasted test subjects, respectively.

In order to generate average AI maps for each feeding condition, individual AI maps from fed and fasted subjects were registered using ANTs [Avants et al., 2011] software.

Once the mean ADC and AI vectors were obtained, we performed statistical analyses to validate the approaches. Fed and fasted pixel populations were created accumulating the mean ADC values of all the pixels from fed or fasted subjects. Briefly, we applied the Wilcoxon rank sum test [Gibbons and Chakraborti, 2011] to investigate statistically significant differences between the medians of the fed and fasted populations. The same procedure was followed for the Fisher Maps approach using the AI vectors.

Finally, we performed a right-tailed paired-sample t-test in each pixel to determine if its AI values in the fasted condition are significantly greater than those in fed condition, a method revealing ultimately which brain regions are most useful to distinguish between fed and fasted conditions.

4.2.2 Results

The performance of Fisher Maps methodology was tested on DWI data from the brain of rodents and humans under a fasting paradigm. A detailed description of the acquisition and pre-processing of the analyzed data can be found in the section 3.1.

Figure 4.2 summarizes the AI results of the complete brain section of mice, as well as in the hypothalamic, thalamic and cortical regions. The upper inset of figure 4.2 shows the precise location of the brain slice investigated (left) and the anatomical structures (right) contained therein (`http://atlas.brain-map.org/#atlas`, May 4th, 2017). Average AI maps of the complete brain section are shown superimposed over the corresponding T_{2w} images under the fed (left) or fasted (right) conditions (figure 4.2 B). A clear shift from blue to red (center color bar) upon fasting can be observed. This reveals that AI increased globally, albeit heterogeneously, through the selected brain slice of fasted mice. More pronounced AI increases (red shifts) are

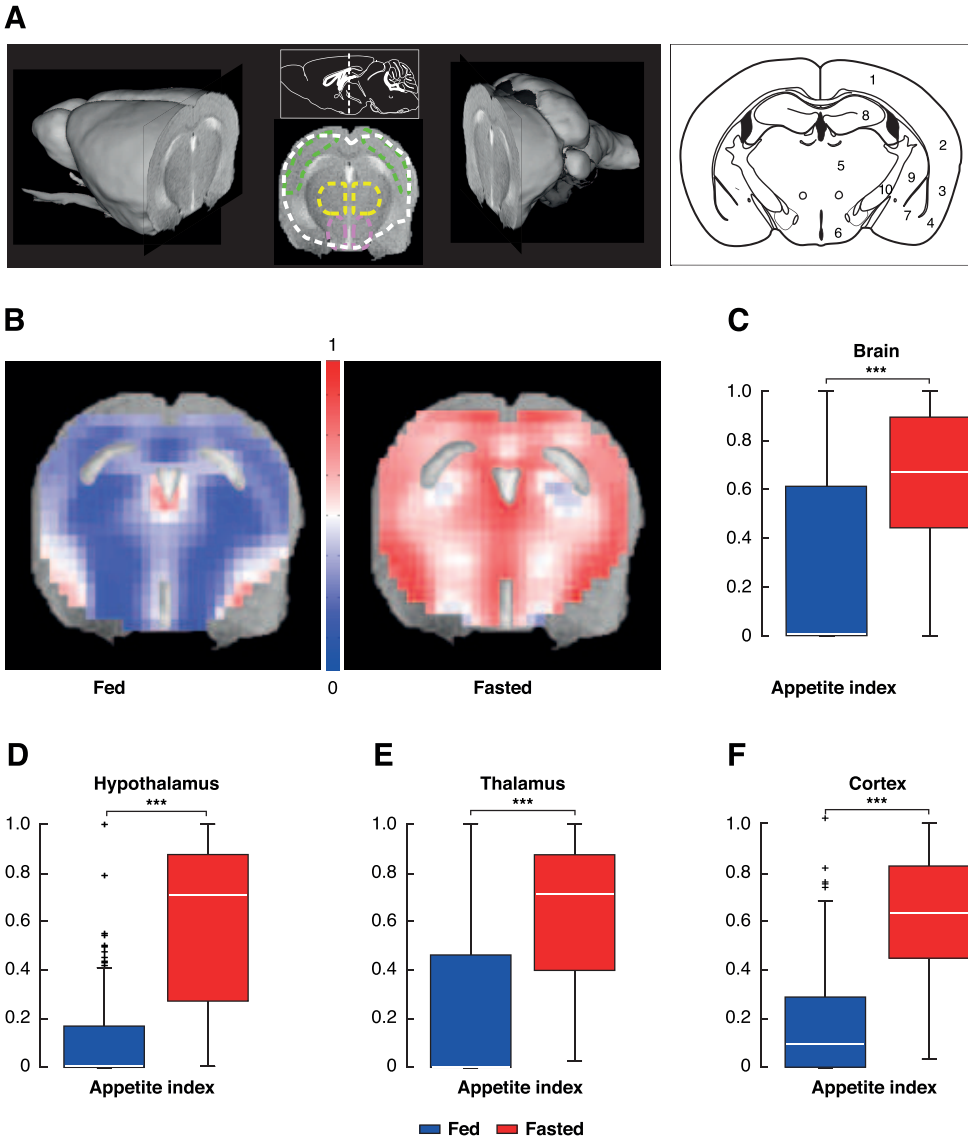

Figure 4.2: Representative AI maps obtained from the selected mouse brain section under the fed or fasted conditions. A: Anatomical location of the slice selected (left) with the ROIs analyzed (inset) and main cerebral structures contained indicated by numbers. B: AI maps from six mice were registered and pixel by pixel mean values calculated. Maps are superimposed over a T_{2w} image of the same slice. Box plots of AI values from the entire brain section (C), the hypothalamus (D), the thalamus (E), and the somatosensory cortex (F) under the fed (blue) or fasted (red) conditions. On each box, the central mark (white line) indicates the median, and the bottom and top edges indicate the 25-th and 75-th percentiles, respectively. The upper and lower limits of the box extend to the most extreme data points not considered outliers, which are plotted individually using the '+' symbol. 1: somatosensory cortex, 2: auditory and temporal association areas, 3: ectorhinal and peririhnal areas, 4: pyriform layers, 5: thalamus, 6: hypothalamus, 7: amygdala, 8: hippocampus, 9: caudoputamen, 10: globus pallidus. (***) p-value < 0.001.

observed in the primary motor and somatosensory cortex including the barrel fields (#1), ventral auditory and temporal association areas (#2), ectorhinal and piriform

areas (#3 and #4), thalamus (#5) and hypothalamus (#6). Smaller increases can also be detected in the amigdala (#7) hypoccampus (#8), caudoputamen (#9), and globus pallidus (#10) regions.

Figure 4.2, from C to F, summarizes the AI changes expressed quantitatively as box plots. In the complete brain section (Figure 4.2 C), median AI values (fed/fasted) increased significantly 0.01/0.72, with similar significant increases in the other regions (Figure 4.2, from D to F). The thalamus showed, however, the largest median appetite stimulation (0.002/0.72) in the mouse brain, with relatively lower appetite stimulation values in the hypothalamus (0.05/0.71) or the motor and somatosensory cortex (0.01/0.63), respectively.

Average AI maps from the brain of seven rats under the fed and fasted conditions are shown (Figure 4.3). The upper panel (figure 4.3 4A) shows the anatomical location of the slice investigated (left) and the main cerebral structures contained (right, https://scalablebrainatlas.incf.org/rat/PLCJB14, May 4th, 2017). Figure 4.3 B provides the average AI maps of the complete brain section while figure 4.3, from C to F, shows box plots of the AI values of the complete brain section, the hypothalamus, the thalamus and the neocortex, respectively.

AI increased heterogeneously in the complete rat brain section as revealed by the shift from blue to red colors, with more pronounced increases (red) in the hippocampal formation (#11), striatum (#12), thalamus (#13), hypothalamus (#14), neocortex (#15) and basal fore brain area (#19). Smaller AI increases can also be observed in the fimbria (#16), the stria terminalis (#17), and the corticofugal pathways (#18). Box plots, indicate that median AI values increased during the fed/fasted transition by 0.33/0.77 in the complete brain section (figure 4.3 C), 0.33/0.89 in the hypothalamus (figure 4.3 D), 0.37/0.92 in the thalamus (figure 4.3 E), and 0.31/0.88 in the cortex (figure 4.3 F). Together, these changes reveal that the thalamus is the cerebral structure becoming more stimulated by appetite in the rat brain, with concomitant activations in the hypothalamus and cortex as well as in other cerebral structures. In particular, the relative pattern of appetite stimulation of the rat brain appears to be different from that found in the mouse brain, with larger relative AI values in the thalamic structures, amygdala and globus pallidus in the rat.

Figure 4.4 summarizes the effects of the fasting in the human brain, by showing average AI maps from six subjects. Figure 4.4 A (upper panel) depicts the location of the slice acquired and the main anatomical regions contained (this slide areas can be consulted online at [Allen-Institute, 2017]), while figure 4.4 B shows median AI maps through this section, figure 4.4, from C to F, shows box plots of the AI changes in the complete brain section, the hypothalamus, the thalamus and the cortex, respectively. AI values increased significantly in the complete brain section (figure 4.4 C), the hypothalamus (figure 4.4 D), the thalamus (figure 4.4 E) and the cortex (precentral and temporal gyrus, figure 4.4 F) as revealed by the shift

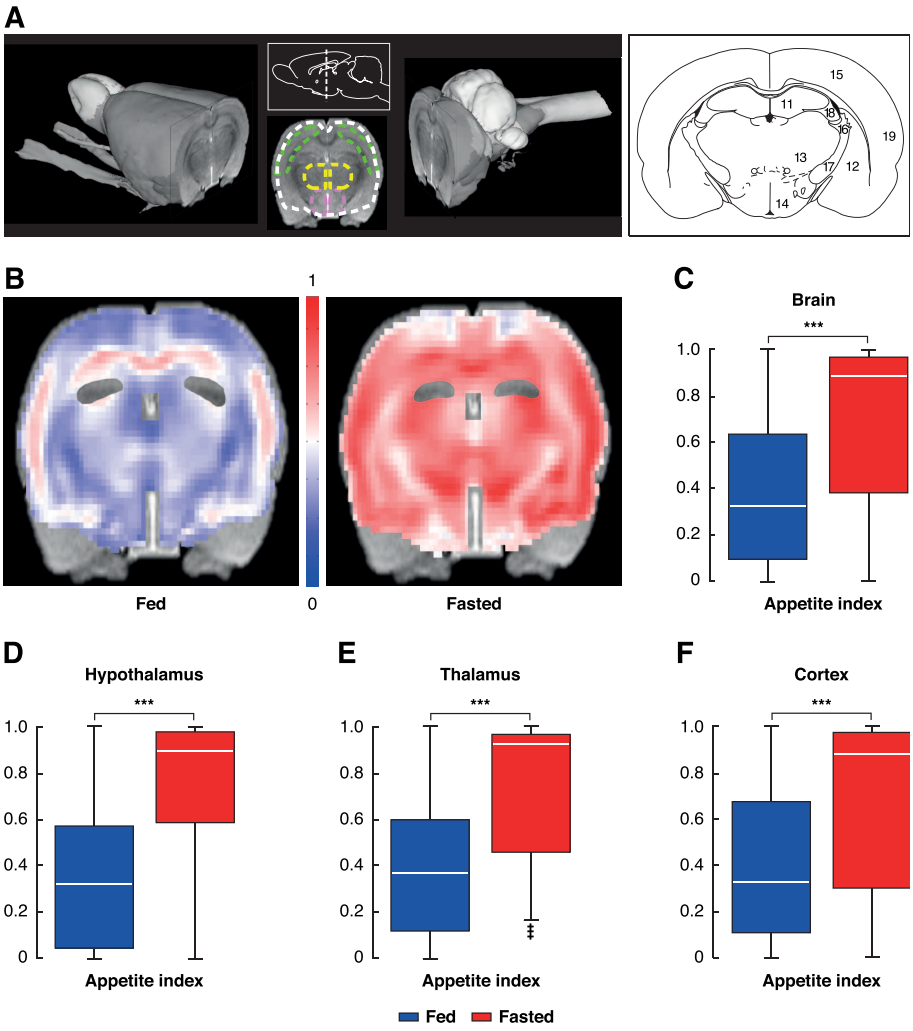

Figure 4.3: Representative AI maps obtained from the selected rat brain section under the fed or fasted conditions. A: Anatomical location of the slice selected (left) with the ROIs analyzed (inset) and main cerebral structures contained (right). B: AI maps from seven animals were registered and pixel by pixel mean values calculated. Calculated AI maps are overlaid over the background of a T_{2w} image. Box plots of AI values in selected ROIs are represented as indicated in Figure 4.2. C: entire brain section, D: hypothalamic region, E: thalamic region, and F: cortical region. 11: hypoccampus, 12: striatum, 13: thalamus, 14: hypothalamus, 15: neocortex, 16-17: stria terminalis and corticofugal pathways, 18: fimbria of the hypoccampus, 19: basal forebrain region. (***) p-value<0.001.

from light blue to red values. In general, appetite scores increased more uniformly in the complete human brain section than in the rodent brain. More specifically, increases in AI can be observed in the precentral and postcentral gyrus (#20), postcentral gyrus (#21), globus pallidus (#22), amygdala (#23), temporal gyrus (#24), putamen (#25), hypothalamus (#26), as well as the white matter of the

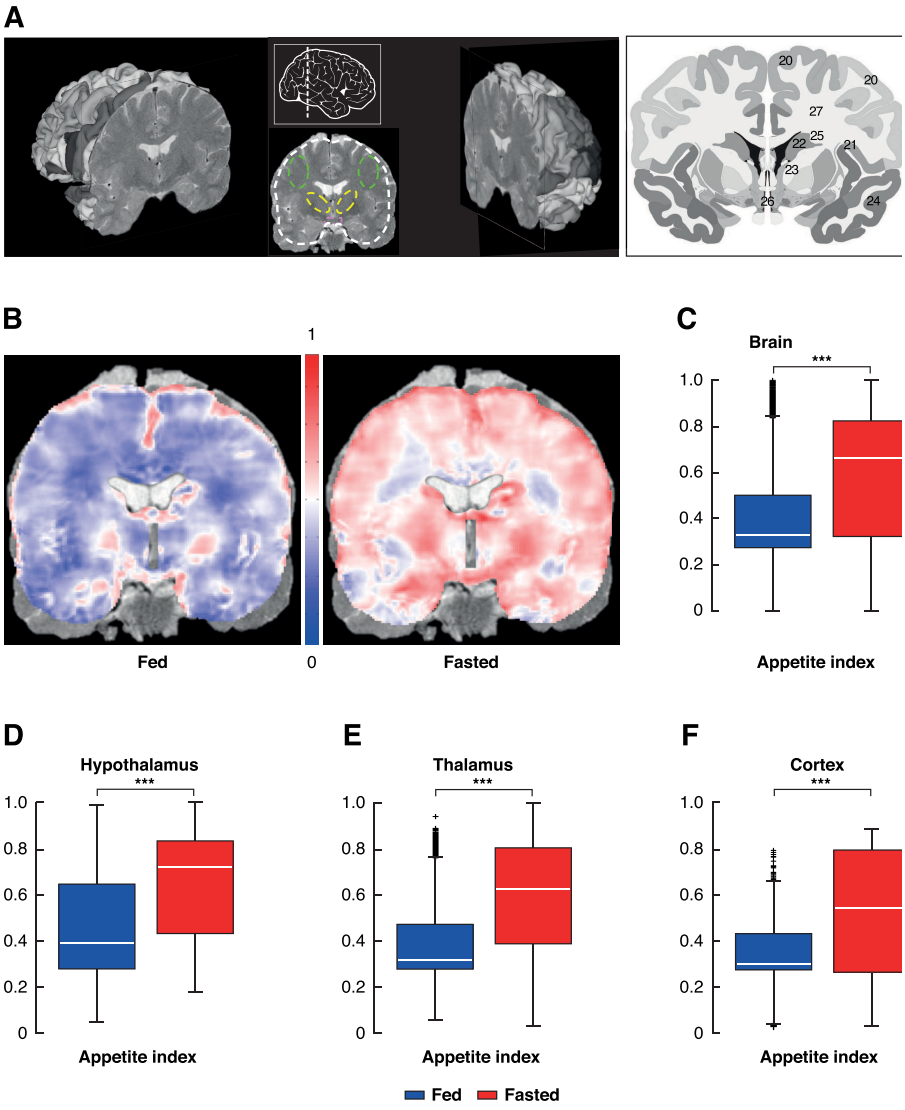

Figure 4.4: Representative AI maps from the selected human brain section under the fed or fasted conditions. A: Anatomical location of the fDWI slice selected (left) with the ROIs analyzed and relevant anatomical structures (right). AI maps were registered and pixel by pixel average values calculated. Maps are overlaid over a T_{2w} image background (B). Box plots of AI values in selected ROIs (C-F) are represented as indicated in Figure 4.2. C: complete brain section. D: hypothalamic region, E: thalamic region, F: cortical region. 20: precentral and postcentral gyrus, 21: parietal operculum and short insular gyri, 22: globus pallidus, 23: amygdaloid complex, 24: superior, median and inferior temporal gyrus, 25: putamen, 26: hypothalamus, 27: white matter of the fore brain. (***) p-value<0.001.

forebrain (#27).

Basal median values of AI in the fed state through the complete human brain section were 0.33, 0.39 in the hypothalamus, and 0.32 or 0.30 in the thalamus or the brain cortex, respectively. AI values increased significantly upon fasting in the complete brain section (0.66), the hypothalamus (0.72), the thalamus (0.63) and the brain cortex (0.55), although to a smaller extent than in the rodent brain, disclosing lower incremental appetite scores in humans.

Summarizing, median AI values under resting or appetite stimulation conditions, revealed considerable differences between the three species with more intense basal stimulations in the satiated brain of rats and humans than in mice, and lower appetite stimulation AI scores in humans than in rodents, respectively.

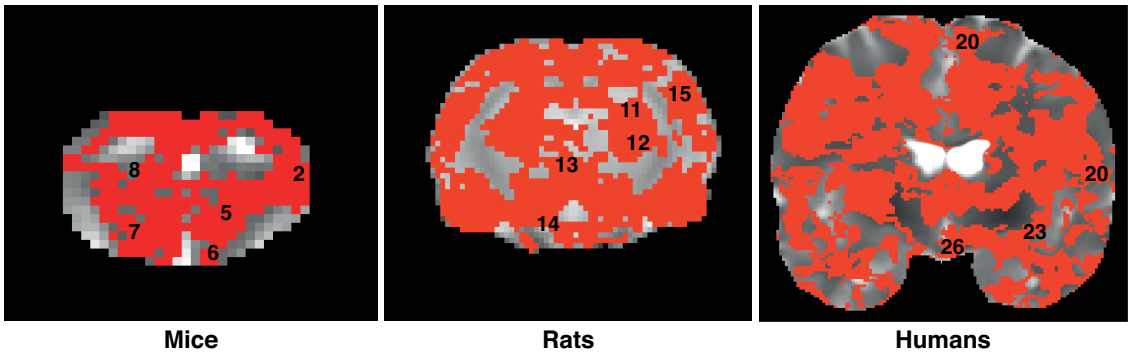

Figure 4.5: Right-tailed paired-sample t-test, applied pixel by pixel, to registered AI maps from mice (left), rats (center) and humans (right). Pixels in red show areas where AI is significantly greater ($\alpha = 0.05$) in the fasted than in the fed condition. The probability maps appear superimposed over a basal DWI of the corresponding reference subjects used to register. Numbers refer to specific regions mentioned in the text.

Figure 4.5 shows the results of applying, pixel by pixel, a right-tailed paired-sample t-test, to the registered AI maps from six mice (left), seven rats (center) and six humans (right), under the fed and fasted conditions, respectively ($\alpha = 0.05$). Pixels in red indicate that the corresponding AI value is significantly higher in the fasted than in the fed condition. Briefly, mice showed significant increases in AI upon fasting through the pixels of the temporoparietal cortex (#2), hippocampus (#8), thalamus (#5), the hypothalamus (#6) and amygdala (#7). Rats also showed significant AI increases in the neocortex (# 15), the hippocampus (#11), the striatum (#12), the thalamus (#13) and the hypothalamus (#14). Taken together, probability maps from rodent models indicate that the median AI values calculated for the different brain regions depicted in figures 4.2 and 4.3, are well reflected by significant pixel by pixel differences. Probability maps in humans presented a more complex appetite activation pattern than in rodents, with multiple significant activation areas. Briefly, the most significant areas activated by appetite in the human brain included, among others, pixels in the pre-central, post-central and temporal gyri (#20), amygdala complex (#23) and hypothalamus (#26).

On the other hand, we fitted the mono-exponential model to the DWI signal attenuation [Alexander et al., 2007] as described by

$$S'_i(b,d) = e^{-b \cdot ADC_i(d)} \quad 1 \leq i \leq n \quad 1 \leq d \leq 3 \tag{4.5}$$

where $ADC_i(d)$ represents the apparent diffusion coefficient of i-th pixel in d-th direction.

The apparent diffusion coefficient was determined for each pixel intensity vector and orthogonal direction in both feeding conditions, using the non-linear least-squares fitting Trust-Region algorithm [Conn et al., 2000]. We used increasing values for goodness of fit (r^2) in the range $0.7 < r^2 < 0.9$, to investigate the number of pixels rejected for increasing fitting accuracies (Table 4.1). We chose $r^2 = 0.8$ as an optimal balance between the accuracy imposed and number of pixels accepted for the fit. For each subject-condition, the ADC value of each pixel in the three diffusion directions was averaged, obtaining a mean ADC vector per subject-condition.

Figure 4.6 summarizes ADC measurements through the selected brain regions in mice (4.6 A), rats (4.6 B) or humans (4.6 C) under the fed (blue) or fasted (red) conditions. The left panels show representative T_{2w} images of the complete brain section, delineating the regions of interest (dashed lines) used for ADC evaluation in the complete brain (white), cortical (green), thalamic (yellow) and hypothalamic (purple) areas. Median ADC values overcoming the $r^2 > 0.8$ fitting criteria are presented as box plots in the corresponding right panels.

In mice, fasting induced a significant decrease in the median ADC ($\mu m^2/s$) values (fed/fasted) of the complete brain section (599/542), as well as in the hypothalamus (600/532), thalamus (585/522) and cortex (578/519). Rats showed also decreased median ADC values in the complete brain section (978/884), hypothalamus (1000/880), thalamus (960/889) and cortex (926/846). Humans depicted a similar response with significant decreases in the ADC of the complete brain (778/756), hypothalamus (941/812) and neocortex (802/751), without appreciable ADC changes in the thalamus (724/713).

Mono-exponential fittings imposed the rejection of considerable number of pixels to accommodate the accuracy restrictions. Table 4.1 shows the total number of pixels used and rejected in every region of interest (ROI), to overcome accuracy limitations in the range $0.7 \leq r^2 \leq 0.9$. For a total number of pixels (fed/fasted) in the entire brain section of mice (2018/1981), rats (12807/13173) or humans (68412/68496) the percentage of pixels overcoming the accuracy restrictions decreased, as expected, for increasing r^2 values. The percentage of pixels overcoming the $r^2 < 0.7$ limitation were (%fed/%fasted): 34/49 in mice, 57/13 in rats and 55/58 in humans, decreasing markedly to 25/44 in mice, 47/12 in rats and 4/4 in humans for the$r^2 < 0.9$ limitation. Similar results were obtained in selected regions as hypothalamus, thalamus and cortex, resulting often in incomplete ADC maps lacking the information from rejected pixels in individual subjects. Together, these results suggested a subop-

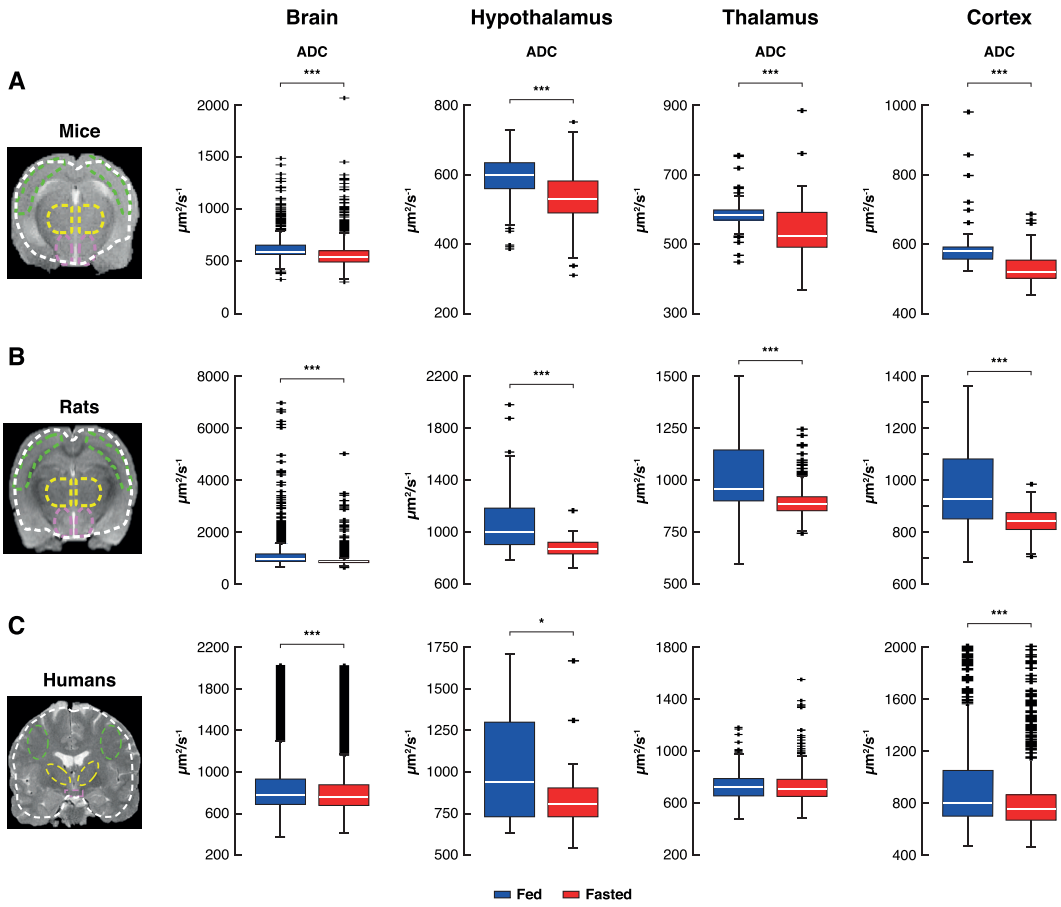

Figure 4.6: Cerebral ADC values in the brain of mice (A), rats (B) and humans (C) under the fed or fasted conditions. Left panels: ROI perimeters selected for analysis are superimposed over representative T_{2w} images of the different models and indicated by the dashed color lines; Complete brain (white), hypothalamus (purple), thalamus (yellow) and cortex (green). Right panels: Box plots of the ADC values from the entire brain section, the somatosensory cortex, the thalamus and the hypothalamus, under the fed (blue) or fasted (red) conditions. (*) p-value<0.05, (***) p-value<0.001.

Experimental System	Selected ROI	Total pixels processed		% Remaining pixels* Fitting Accuracy					
				$r^2 > 0.7$		$r^2 > 0.8$		$r^2 > 0.9$	
		Fed	Fasted	Fed	Fasted	Fed	Fasted	Fed	Fasted
Mice	Whole Brain	2018	1981	34	49	32	48	25	44
	Hypothalamus	232	223	32	52	31	52	23	48
	Thalamus	253	253	33	50	33	49	32	48
	Cortex	198	149	38	54	37	54	34	54
Rats	Whole Brain	12807	13173	57	13	56	13	47	12
	Hypothalamus	630	630	56	14	55	14	42	14
	Thalamus	1890	1875	57	14	57	14	53	14
	Cortex	1680	1680	57	14	56	14	48	14
Humans	Whole Brain	68412	68496	40	44	21	24	4	4
	Hypothalamus	411	408	16	26	7	9	1	1
	Thalamus	3924	3924	27	35	11	16	1	2
	Cortex	2773	2787	40	45	18	24	4	4

Table 4.1: Pixels overcoming increasing accuracy restrictions in the mono-exponential diffusion model. (*)Percentage of pixels overcoming increasing accuracy restrictions $0.7 \leq r^2 \leq 0.9$

timal description of the experimental data by the mono-exponential water diffusion model. Previous bi-exponential model fittings involved similar limitations [Lizarbe et al., 2013], indicating that the complex DWI signal changes observed in the brain upon fasting may not be adequately described in terms of either the mono- or bi-exponential models of diffusion.

4.3 Image classification

4.3.1 Methods

To evaluate image datasets derived from a small number of subjects, we propose a supervised setting where each pixel from the images is an observation, and the intensities values along the acquired images are the features. The image dataset is represented then by f features measured on m observations, each one belonging to one of the two classes investigated: labeled “fed” or “fasted”, “low-grade” or “high-grade”, “female” or “male”. For example, in the appetite study of six mice fasting 16h, $m = 12xr$ and $f = sxd = 18$, where r is the number of pixels from brain, $s = 6$ (the number of b-values), and $d = 3$ (the number of orthogonal diffusion directions) investigated. Figure 4.7 illustrates the two classification methods implemented using this supervised setting: Pixel Ensemble-Based Classification (PE-BC) and Pixel-Based Classification (PB-C), respectively.

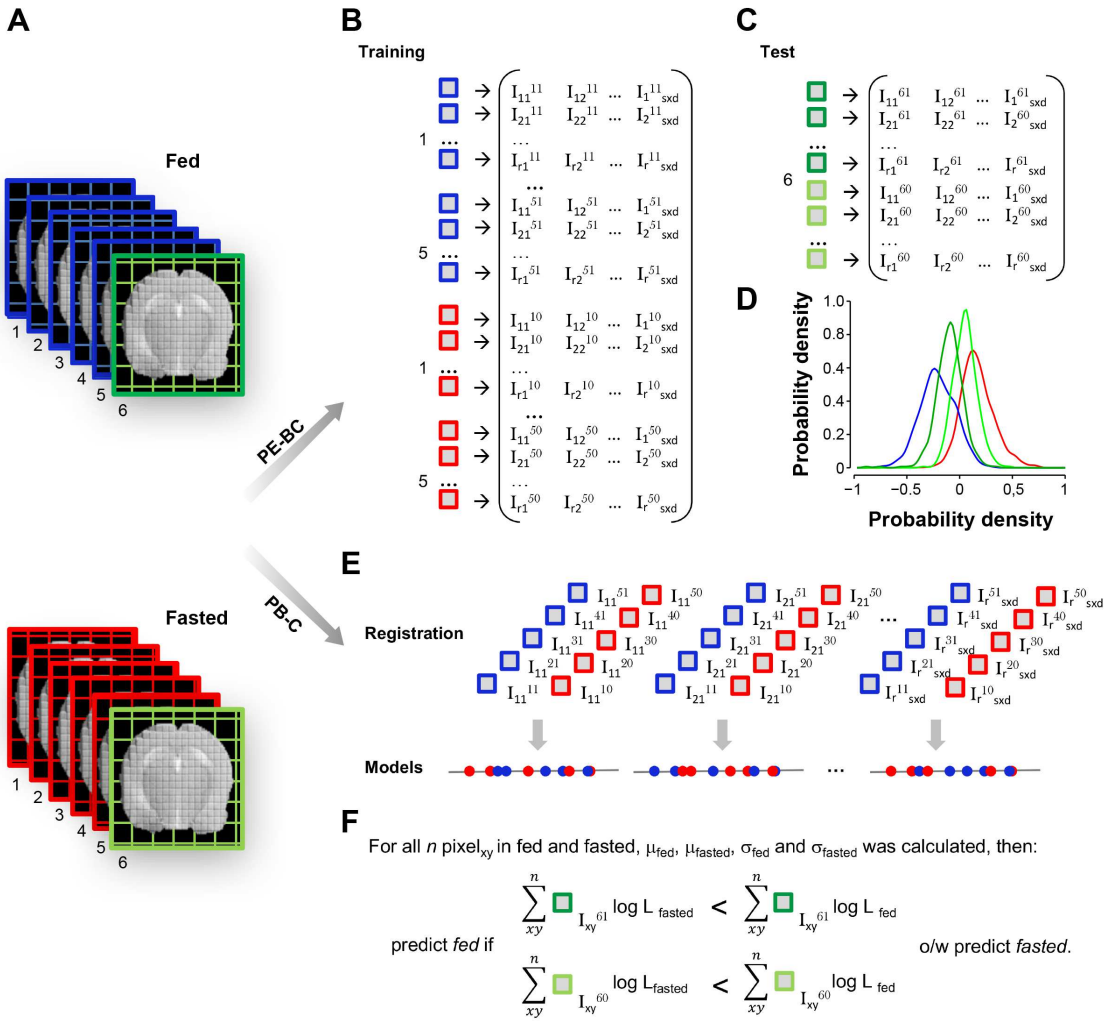

Figure 4.7: Overview of the PE-BC and PB-C classification approaches. A: DWI dataset from animals in the fed ($n = 6$, blue) or fasted ($n = 6$, red) states, after preprocessing. The slices in green represent those belonging to the test subject in this example. B, C and D panels correspond to the Pixel Ensemble-Based Classification (PE-BC) method. B: Matrix representation of the combined DWI pixel collections from the fasted and fed subjects, forming the training set. C: Matrix representation of the combined DWI pixel collection from the test subject. D: Visual representation of the PE-BC classification method. E and F panels correspond to the P-BC method. E: Registered pixels from the image dataset. For each pixel, a model is calculated using its intensities values in the images from training set. F: The predicted class for the subject test is obtained by comparing the sum of log-likelihood$_{\text{fed}}$ and the sum of log-likelihood$_{\text{fasted}}$ for all pixels in test set. If the sum of log-likelihood$_{\text{fed}}$ is bigger than the sum of log-likelihood$_{\text{fasted}}$, the predicted class is fed, otherwise is fasted.

The PE-BC method (Figure 4.7, B, C and D) is the extension of the Fisher maps approach described above (Section 4.2), adding a classification criterion to predict the class of a subject based on the class assigned to the pixels in their images. Our model consists of two gaussian distributions, one for each class. The gaussian distribution of each class is fitted using the mean and standard deviation of the

Fisher projection (vector z in equation 4.1) computed for each class. To classify a new subject, we compute the log-likellyhood of gaussian distribution against each distribution in the model, getting $\text{loglikelihood}_{fed}$ and $\text{loglikelihood}_{fasted}$. Finally, the predicted class (C_{pred}) is obtained using the rule

$$C_{pred} = \begin{cases} \text{fed} & \text{if} \quad \text{loglikelihood}_{fed} > \text{loglikelihood}_{fasted} \\ \text{fasted} & \text{if} \quad \text{loglikelihood}_{fed} \leq \text{loglikelihood}_{fasted} \end{cases} \tag{4.6}$$

Moreover, in the P-BC approach, all images are registered to allow pixel by pixel comparisons. For each pixel, a model is obtained applying Naïve Bayes [Domingos and Pazzani, 1997] to its intensities values in the images from both classes. Mean and standard deviation in both classes were calculated also for all pixels. Then, we computed the sum of $\text{log-likelihood}_{\text{fed}}$ and the sum of $\text{log-likelihood}_{\text{fasted}}$ for all pixels. Lastly, to classify a new subject, we used the rule

$$C_{pred} = \begin{cases} \text{fed} & \text{if} \quad \sum_{xy}^{r} \text{loglikelihood}_{fed}^{xy} > \sum_{xy}^{r} \text{loglikelihood}_{fasted}^{xy} \\ \text{fasted} & \text{if} \quad \sum_{xy}^{r} \text{loglikelihood}_{fed}^{xy} \leq \sum_{xy}^{r} \text{loglikelihood}_{fasted}^{xy} \end{cases} \tag{4.7}$$

where x and y represent the coordinates of the pixels.

As mentioned in Section 4.2.1, we implemented a leave-one-subject-out strategy. The classifier is trained, therefore, with the images of all the subjects except one, which was used as test. We repeated the same process using in each case a different subject as test, generating different data partitions. The test error in the classification is averaged across the partitions.

4.3.2 Results

Pixel Ensemble-based classification

We investigated the performance of PE-BC method in the classification of MR images of the datasets from studies of appetite stimulation and glioma grading. A detailed description of the acquisition and pre-processing of the analyzed data can be found in the sections 3.1 and 3.2. Using a leave-one-subject-out strategy, we obtained the classification accuracies indicated in tables 4.2 and 4.3.

In general, the classification accuracies were greater than 74%. In appetite stimulation datasets, the highest accuracy was 83% classifying twelve conditions (fed or fasted) in six investigated human volunteers. In gliomas studies datasets, the largest accuracy was 92% classifying twelve high-grade glioma rat models into male and female.

Dataset*	Description	Subjects number	Features number	Accuracy
Aptetite_1	Mouse model 48h of fasting	8 in two conditions	33 DWI	81
Aptetite_2	Mouse model 16h of fasting	6 in two conditions	18 DWI	75
Aptetite_3	Rat model 16h of fasting	7 in two conditions	9 DWI	79
Aptetite_4	Human model 24h of fasting	6 in two conditions	21 DWI	83

Table 4.2: Classification results in appetite stimulus datasets. (*)Dataset identifier. Appetite_1: eight male mice investigated under fed and 48h after fasting. Appetite_2: six male mice investigated under fed and 16h after fasting. Appetite_3: seven male rats investigated under fed and 16h after fasting. Appetite_4: six human volunteers investigated under fed and 24h after fasting.

Dataset*	Description	Subjects number	Features number	Accuracy
Tumor_1	Rat model Low/high-grade glioma	10 low-grade 10 high-grade	1 MT 1 T2w	90
Tumor_2	Human patients Low/high-grade glioma	17 low-grade 18 high-grade	1 FLAIR post gad 1 T1w pre gad	83
Tumor_3	Rat model Gender differences in a high-grade glioma	6 male 6 female	14 DTI	92

Table 4.3: Classification results in brain tumors datasets. (*)Dataset identifier. Tumor_1: low and high-grade glioma rat model. Tumor_2: low and high-grade glioma in human patients. Tumor_3: gender differences in a high-grade glioma rat model.

Pixel-Based Classification

We applied the P-BC method in the appetite datasets (Table 4.4). The highest accuracy was 67% classifying twelve conditions (fed or fasted) in six investigated mice or human volunteers. This method was not applied in the datasets from glioma studies since it is not appropriate for this problem, because it makes pixel-to-pixel comparisons.

Validation strategy

It is necessary to prove, however, that the results of the classification using PE-BC are not derived from an arbitrary coincidence. To do this, we implemented a strategy based on nonparametric permutation test [Fisher, 1935, Bullmore et al., 1999, Nichols and Holmes, 2002]. This strategy consisted of four steps: (i) to gener-

Dataset*	Description	Subjects number	Features number	Accuracy
Appetite_1	Mouse model 48h of fasting	8 in two coditions	1 DWI	67
Appetite_2	Mouse model 16h of fasting	6 in two conditions	1 DWI	58
Appetite_3	Rat model 16h of fasting	7 in two conditions	1 DWI	64
Appetite_4	Human model 24h of fasting	6 in two conditions	1 DWI	67

Table 4.4: Classification results in appetite stimulus datasets. (*)Dataset identifier. Appetite_1: eight male mice investigated under fed and 48h after fasting. Appetite_2: six male mice investigated under fed and 16h after fasting. Appetite_3: seven male rats investigated under fed and 16h after fasting. Appetite_4: six human volunteers investigated under fed and 24h after fasting.

ate all possible artificial training datasets by randomly reassigning the class labels to each subject or subject-condition; (ii) to obtain the Fisher projection for both classes in each artificial training dataset; (iii) to measure the class separation (s) for each artificial training dataset and, (iv) to compare this separation with the separation obtained for the real classes. The class separation is given by the expression:

$$S = \frac{(\mu_{\text{classA}} - \mu_{\text{calssB}})^2}{\sigma^2_{\text{classA}} + \sigma^2_{\text{classB}}} \tag{4.8}$$

where, μ_{classA} and μ_{classB} represent the averages of the Fisher projection for each class, and σ_{classA} and σ_{classB} represent the standard deviation of the Fisher projection for each class.

Two different protocols were implemented to generate the artificial training datasets: First, we selected a subset of all the subjects, and for every selected subject, we exchanged the class labels of its two subject-conditions (one for each class). For example, for Appetite_1 dataset, there are 2^8 possible subsets made with 8 mice, but only half of these must be considered because of the symmetry in class exchange. So a total of $2^7 = 128$ different artificial training datasets are obtained. Second, we randomly reassigned each subject-condition to one of the two class labels. For Appetite_1 dataset, the number of possible artificial training datasets for 16 subject-conditions (two for each mouse) is 2^{16}, but only half must be considered, as discussed above, to account for the symmetry in class exchange. Thus, a total of $2^{15} = 32768$ different artificial training datasets are generated in this case.

For each of the two protocols, we applied LDA and obtained the Fisher projection for each one of the generated artificial training datasets; comparing the class separation with those obtained using the real classes. We hypothesize that; if the linear projection obtained by LDA is capturing a property related to appetite or tumor, rather than statistical fluctuations, then the degree of separation between

the projected classes in the original training dataset will be amongst the best that can be obtained arbitrarily.

Figures 4.8, 4.9, 4.10, and 4.11 show the separability normalized histogram for all artificial training datasets that can be constructed using the first (A panels) and second (B panels) protocols in the appetite stimulation datasets. Additionally, figure 4.12 illustrates the separability normalized histogram for all artificial training datasets obtained using the second protocol in the glioma datasets. It is important to note that, in most cases, the real training dataset is in the top 10% separability. The results of this validation strategy have been less good in the Appetite_2 and Tumor_3 datasets, in which 78% and 70% of the artificial training datasets, depict worse separability than the real training dataset, respectively.

4.4 Discussion

The methodologies presented in this chapter illustrate well the kind of support that supervised learning methods can provide for non-invasive, functional studies of cerebral performance, and for glioma grading and characterization.

4.4.1 Fisher maps

Interpretation of Apparent Diffusion Coefficient and Appetite Index

ADC and AI values of cerebral activation by appetite represent different types of information. Although both are generated from the same DWI datasets and include initially the same diffusion information, the DWI datasets originating ADC and AI values are processed differently. ADC values indicate the cerebral distribution of the mono-exponential apparent diffusion coefficient of water, whereas AI values indicate the probability that a given pixel belongs to the fasted condition. Thus, while ADC changes are classically interpreted in phenomenological terms, as revealing changes in intracellular and extracellular water diffusion [Le Bihan, 2012, Norris, 2001, Syková and Nicholson, 2008, Ackerman and Neil, 2010], AI values may be interpreted directly as the relative appetite activation (or "appetite score") of a pixel, in a relative scale of 0 (fed) to 1 (fasted). In this sense, both ADC and AI values include the contribution of all physiological processes underlying water diffusion [Le Bihan, 2007, Le Bihan and Iima, 2015]. However, while the precise mechanistic interpretation of ADC changes remains uncertain [Le Bihan and Iima, 2015], AI expresses the appetite activation directly as an "appetite score", overcoming the mechanistic interpretations. Since both strategies are not mutually exclusive and can be applied to the same dataset, AI maps may provide a useful complement

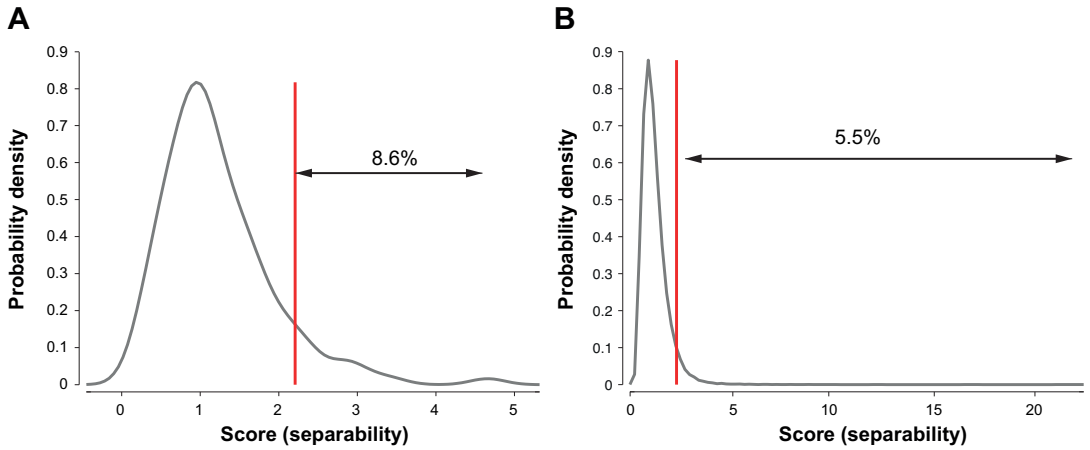

Figure 4.8: Normalized histograms of class separation for each of the artificial appetite training datasets, as generated by randomly reassigning the class labels in Appetite_1 and using the PE-BC methodology. Appetite_1 dataset: eight male mice investigated under fed and 48h after fasting. The vertical red line corresponds to the separability for the real training dataset. A: Artificial training datasets obtained exchanging the class labels of images for a given subject. B: Artificial training datasets generated for all possible class label reassignments. The real training dataset is in the top 8.6%/5.5% for the first/second test.

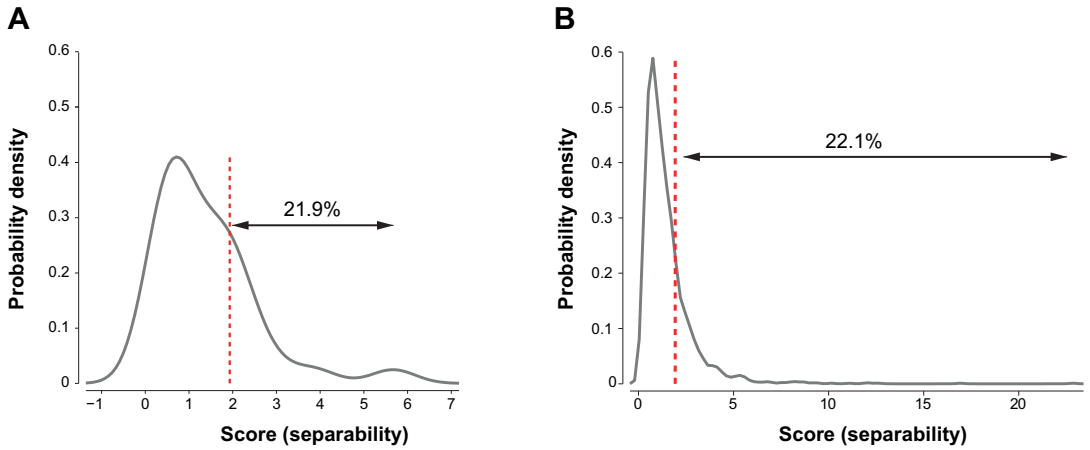

Figure 4.9: Normalized histograms of class separation for each of the artificial appetite training datasets, as generated by randomly reassigning the class labels in Appetite_2 and using the PE-BC methodology. Appetite_2 dataset: six male mice investigated under fed and 16h after fasting. The vertical red line corresponds to the separability for the real training dataset. A: Artificial training datasets obtained exchanging the class labels of images for a given subject. B: Artificial training datasets generated for all possible class label reassignments. The real training dataset is in the top 21.9%/22.1% for the first/second test.

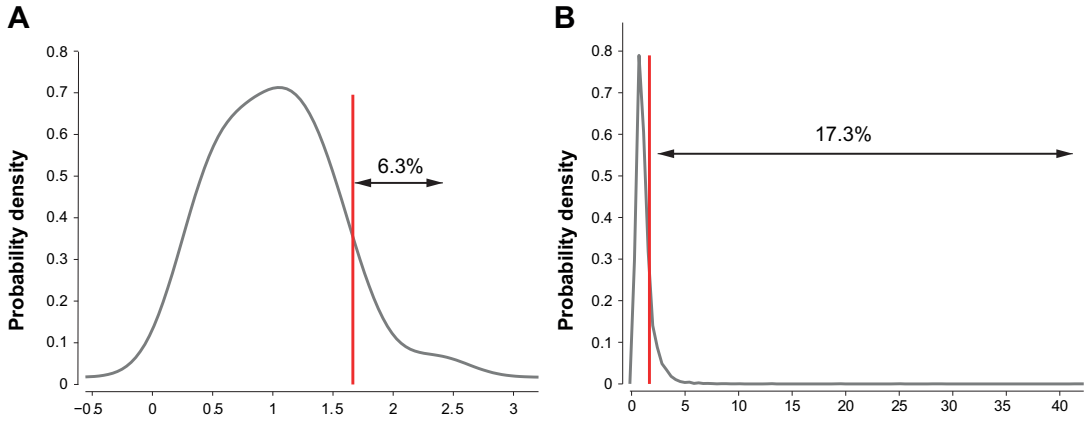

Figure 4.10: Normalized histograms of class separation for each of the artificial appetite training datasets, as generated by randomly reassigning the class labels in Appetite_3 and using the PE-BC methodology. Appetite_3 dataset: seven male rats investigated under fed and 16h after fasting. The vertical red line corresponds to the separability for the real training dataset. A: Artificial training datasets obtained exchanging the class labels of images for a given subject. B: Artificial training datasets generated for all possible class label reassignments. The real training dataset is in the top 6.3%/17.3% for the first/second test.

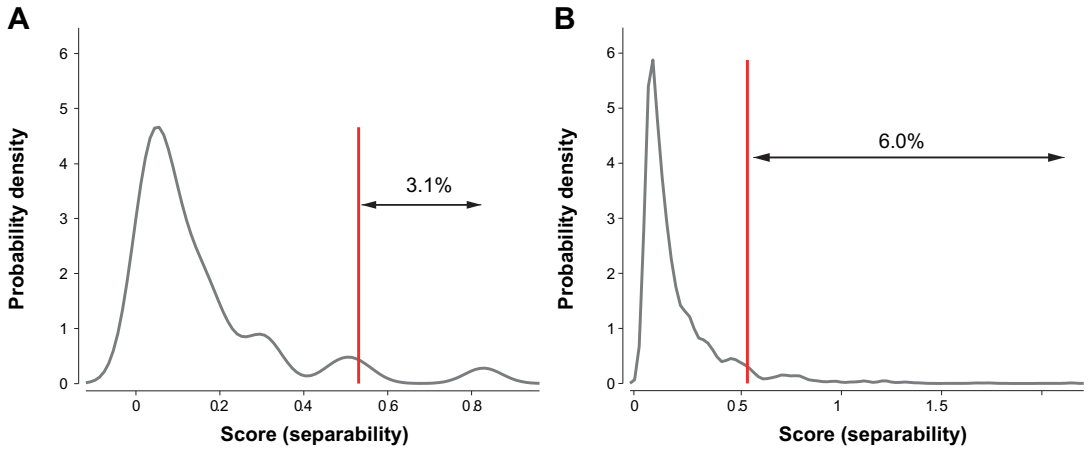

Figure 4.11: Normalized histograms of class separation for each of the artificial appetite training datasets, as generated by randomly reassigning the class labels in Appetite_4 and using the PE-BC methodology. Appetite_4 dataset: six human volunteers investigated under fed and 24h after fasting. The vertical red line corresponds to the separability for the real training dataset. A: Artificial training datasets obtained exchanging the class labels of images for a given subject. B: Artificial training datasets generated for all possible class label reassignments. The real training dataset is in the top 3.1%/6% for the first/second test.

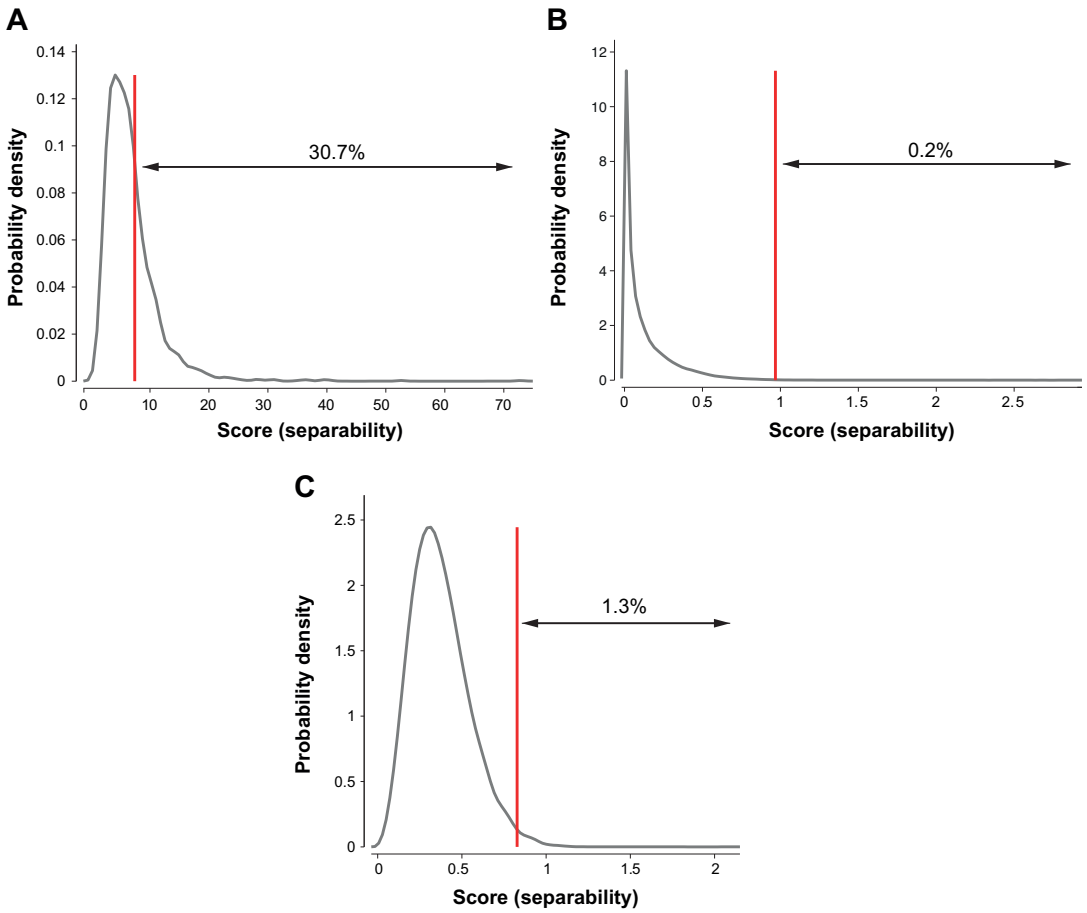

Figure 4.12: Normalized histograms of class separation for each of the artificial tumor training datasets, as generated by randomly reassigning the class labels in the PE-BC methodology. A, B and C: Artificial training datasets generated doing all possible class label reassignments. The vertical red line corresponds to the separability for the real training dataset. A: Tumor_1 dataset, low and high-grade glioma rat model. The real training dataset is in the top 30.7%. B: Tumor_2 dataset, low and high-grade glioma in human patients. The real training dataset is in the top 0.2%. C: Tumor_3 dataset, gender differences in a high-grade glioma rat model. The real training dataset is in the top 1.3%.

that allows to relate the results of the classical mechanistic interpretations, to the resulting fasting activation scores, thus providing independent criteria to evaluate the results of classical statistical parametric mapping strategies [Ashburner, 2012].

Apparent Diffusion Coefficient

Fasting induced global and heterogeneous decreases in brain ADCs, suggesting that cerebral activation by appetite occurs simultaneously, albeit to different extents, through the complete brain section of rodents and humans. Decreases in cerebral ADC have been previously reported in ischemic episodes and assigned to neurocellular swelling events, although the precise relationship between the ADC changes and the swelling response remains elusive [Lo et al., 1997]. Interestingly, the patterns of activation appeared to be different in rodents and humans, with cortical and hypothalamic regions showing the largest ADC reductions in all cases and the cortico-limbic system depicting more pronounced ADC reductions in rodents than in humans. This suggests that a more intense activation of the hedonic areas of the human brain may demand, additionally, the presentation of food cues or exogenous appetite related peptides [De Silva et al., 2012, Batterham et al., 2007].

Appetite Index

AI maps identified automatically the cerebral homeostatic and hedonic centers stimulated by appetite. Notably, AI maps provided directly relative appetite scores for every region, favoring the establishment of fasting activation hierarchies within the different cerebral structures visualized. As in the case of ADC, the priorities of AI changes were different in the rodent models or in humans. While hypothalamic activation was detected in all cases, rodents showed more pronounced relative activations of the cortico-limbic areas than humans, respectively. Moreover, the integral appetite circuitry analysis provided by AI maps, revealed appetite activation in brain regions not previously associated to processing of the appetite stimulus [Fulton, 2009], as the temporoparietal or auditory cortex and pyriform areas in mice; the hippocampal fimbria, the stria terminalis, corticofugal pathways and basal fore brain area in rats; or the cingulate gyrus in humans.

Comparisons of Apparent Diffusion Coefficient and Appetite Index

In general, the results from AI maps, matched well those obtained with the mono-exponential fittings, validating mutually both approaches. Briefly, both showed simultaneous activation of hypothalamic and cortico-limbic structures, albeit with different profiles in rodents and humans. However, AI maps revealed appetite changes

within a larger dynamic range than that obtained with the ADC maps. While the AI values varied in a full range from 0 to 1, the ADC decreases observed after fasting did not exceed more than 10% of the fed values in all cases. Together, these results complement previous MEMRI approaches in the rodent brain [Kuo et al., 2006], or PET and fMRI BOLD studies in the human brain [Batterham et al., 2007, Atalayer et al., 2014, Ely et al., 2014, Geliebter et al., 2013], demonstrating functional responses in homeostatic and hedonic centers after the administration of exogenous peptides involved in appetite regulation or the exposure to different types of food (or non-food) cues. However, the present study contributes additionally, comparative interspecies differences that may become important during the translational evaluation of anti-obesity or anti-diabetic therapies, in animal models and humans.

4.4.2 Image classification

We report two classification methods and a new supervised setting addressing the MRI classification problem, using the pixel collection dataset rather than the image collection dataset employed classically. The best classification results were obtained using PE-BC method. PE-BC method, based on LDA, is easy to implement, quick, robust, and only requires two pre-processing steps, slice and ROI selection and intensity normalization. Notwithstanding, depending on the problem being analyzed, some additional preprocessing operations may be required.

Important limitations of most MRI analysis protocols include the need of numerous pre-processing steps as registration, noise reduction, normalization, region of interest selection and dimensionality reduction, all of them involving different requirements and objectives. These pre-processing procedures are based on extensive and complex mathematical treatments involving important assumptions, frequently leading to significant approximations to be able to describe the real problem in terms of the theoretical frame. Excessive transformations may, in addition, degenerate the information contained in the image and thus, minimizing the number of pre-processing steps seems recommendable when implementing any image processing protocol.

The proposed methodologies may improve this situation. In particular, the PE-BC method does not require performing registration, as in the conventional setting, remaining less costly in terms of computing time, and avoiding the limitations of this operation. The registration operation constitutes, precisely, a fundamental problem in biomedical imaging and many methods and automatic software have been proposed to perform it [Hill et al., 2001]. Therefore, the selection of one specific registration method to be applied becomes often a difficult decision. Notably, some alignment algorithms incorporate transformations that can modify the intensity value of the image pixels, which may not be appropriate for analyses based on pixel intensities, since these transformations can deteriorate, or invalidate, other-

wise useful information [Hutton and Braun, 2003, Razlighi et al., 2013, Johnson and Fischer, 2015, Guyader et al., 2015]. On these grounds, avoiding the use of registration steps, especially in cases with small number of subjects, represents a valuable strength of our method to maintain all available information for analysis.

We illustrate the performance PE-BC approach by analyzing MRI dataset of cerebral activation in three species subjected to a feeding/fasting paradigm, and MRI datasets of gliomas studies in rodents and humans. In these problems, using a leave-one-out strategy, our classification method obtained valuable results, reaching successes greater than 74% in the appetite stimulation datasets and 82% in gliomas studies datasets. Moreover, the P-BC method, which requires image registration, did not reach accuracies greater than 67% in appetite stimulation datasets.

However, the classification methodology described may be subjected to overfitting. To investigate if overfitting occurred in our experiments, we implemented, besides the leave-one-out strategy, a validation strategy based on applying the classification methodology to artificially constructed class labels, generated from randomly assigned class labels. Since the separability obtained for the real classes is in most cases in the top 10% (considering all possible class label reassignments in all analyzed datasets), we assume there is no evidence of overfitting. Only in two databases, Appetite_2 and Tumor_3, we obtained less good results, probably reflecting noisier images or the contribution of specific tumor heterogeneities, respectively.

In addition, two parameters of the PE-BC method may influence the classification; the number of eigenvectors used in PCA and the elimination of a percentage of the overlapping area within the model distributions. Thus, the PE-BC method presents the limitation of how to select the most appropriate values for these parameters for every specific problem. We suggest that the selection of these parameters might be addressed using cross-validation for large datasets.

In summary, this chapter provided new supervised learning methods to analyze biomedical MR images based on pixel by pixel analysis. Our methods are shown to be able to identify automatically the cerebral structures involved in appetite regulation and to classify glioma grading and sexual dimorphism. These findings provide a measure of confidence supporting further developments in this field.

Chapter 5

Unsupervised learning

In this chapter, we propose a new unsupervised learning algorithm for the analysis of MRI datasets. The first sections describe the mathematical fundamentals and the steps of the algorithm. The following sections describe the results of applying this algorithm to synthetic datasets and to real MRI datasets, as well as the related discussion.

5.1 Introduction

In the biomedical research field, when a pathology is studied, it becomes of outmost interest to identify the patterns shared by all the subjects, referred to as invariant patterns, and also the patterns specific to specific subsets of the investigated population, referred to as specific patterns. Analysis of the invariant and specific patterns can help enormously in the diagnosis, prognosis and selection of the appropriate treatments. Such is the case of primary brain tumors of the central nervous system, named gliomas. These tumors have common features that allow them to be diagnosed and classified in low and high grade, with very different recommended treatments. However, glioma tumours are not homogeneous among patients and at the intratumoral voxel level, justifying why within similar cohorts of patients, some respond favorably to treatment while others do not [Patel et al., 2014, Ke et al., 2014, Victoria et al., 2015].

Figure 5.1 shows the scatter plots for normalized intensities values of tumor pixels in MRI dataset acquired from low (rat 1 to rat 10) and high-grade (rat 11 to rat 20) glioma rat model. The axes of each graph represent the intensity values of each tumor pixel in the two acquired images, MT and T_{2w} images. Observing these graphs it is not possible to find a structure in the data that allows to identify common patterns (pixels clusters, in this case) for all high-grade tumors and for all low-grade tumors; due to the huge variability that exists between the different individuals studied.

Clustering algorithms can be used to find natural patterns in datasets that are beyond human perception. Usually, when these techniques are applied to biomedical images each the data acquired from each subject is analyzed independently, and then, the general structures are extracted. This approach does not consider

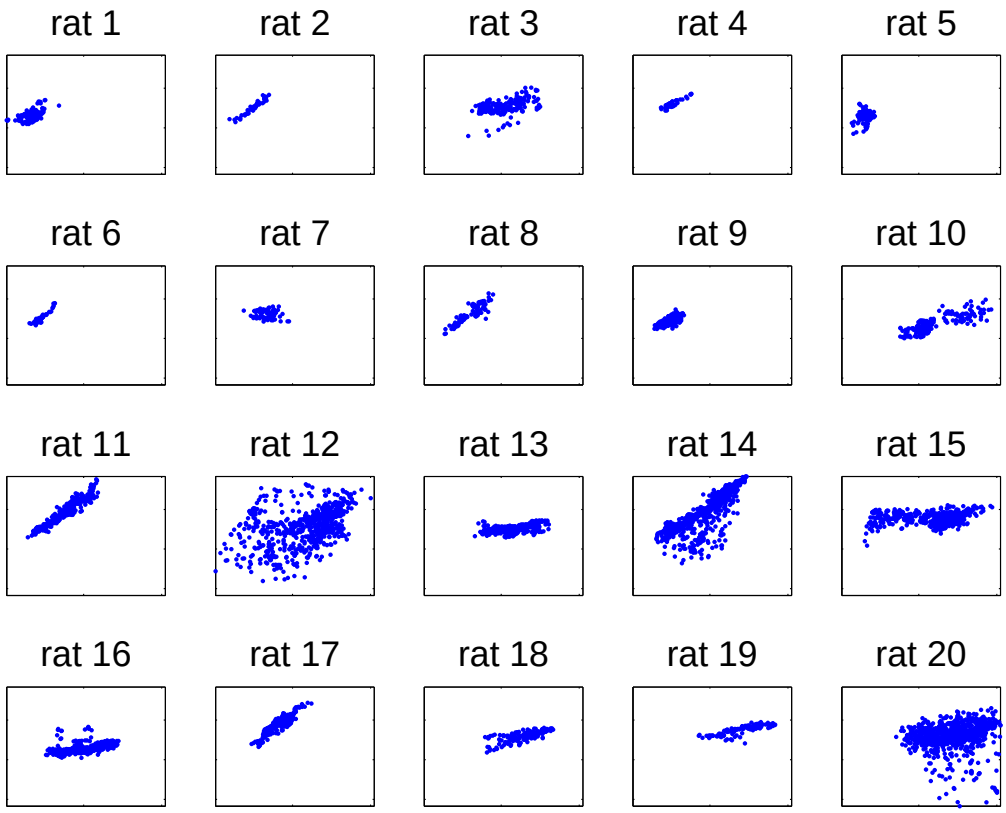

Figure 5.1: Data variability in glioma grade experiments after normalization. Representation of intensities values of tumor pixels in MRI dataset acquired from low (rat 1 to rat 10) and high-grade (rat 11 to rat 20) glioma rat model. The x-axis represents the intensity value of a tumor pixel in the MT image and the y-axis represents the intensity value of the same tumor pixel in the T_{2w} image.

explicitly the relationship between the different investigated subjects, and therefore, shared patterns may be missed. On the other hand, different images from each subject may be combined into a unique dataset, before a standard clustering algorithm is applied, neglecting the different sources of the data and the specific pattern variations associated, precluding to exploit all the available information.

On these grounds, the unsupervised learning method presented in this chapter is a novel clustering algorithm for the joint analysis of datasets obtained from different subjects, detecting both invariant and variant patterns of the data, that are not possible to identify using the standard clustering methods. This idea was inspired in earlier approaches [Pérez and Sánchez-Montañés, 2007, Pérez and Sánchez-Montañés, 2009, Sánchez-Montañés and Baroni, 2010]. We illustrate the performance of this approach using synthetic datasets and MRI datasets of gliomas from animal models.

5.2 The Expectation-Maximization (EM) generic algorithm

The Expectation-Maximization (EM) algorithm [Dempster et al., 1977, Bilmes et al., 1998] is a strategy for parameter estimation in parametric probabilistic models with hidden variables. Let us assume that we have a set of observed data $\mathcal{D}_o$. According to the Maximum Likelihood (ML) parameter estimation strategy, the optimal values for the parameters $\boldsymbol{\Theta}$ of the probabilistic model $\tilde{p}(\mathcal{D}_o|\boldsymbol{\Theta})$ are

$$\hat{\boldsymbol{\Theta}} = \arg\max_{\boldsymbol{\Theta}} \mathcal{L}(\boldsymbol{\Theta}) \tag{5.1}$$

with $\mathcal{L}(\boldsymbol{\Theta})$ being the likelihood

$$\mathcal{L}(\boldsymbol{\Theta}; \mathcal{D}_o) \equiv \tilde{p}(\mathcal{D}_o|\boldsymbol{\Theta}) \tag{5.2}$$

The hidden (unobserved) variables are taken into account using the law of total probability

$$\mathcal{L}(\boldsymbol{\Theta}; \mathcal{D}_o) \equiv \tilde{p}(\mathcal{D}_o|\boldsymbol{\Theta}) = \sum_{\mathcal{D}_h} \tilde{p}(\mathcal{D}_o, \mathcal{D}_h|\boldsymbol{\Theta}) \tag{5.3}$$

Therefore, the ML estimator is determined by equations 5.1 and 5.3.

Since the logarithm is a monotonically increasing function the ML estimator equations can be rewritten in an alternative way as

$$\hat{\boldsymbol{\Theta}} = \arg\max_{\boldsymbol{\Theta}} \ell(\boldsymbol{\Theta}) \tag{5.4}$$

with

$$\ell(\boldsymbol{\Theta}; \mathcal{D}_o) \equiv \ln \tilde{p}(\mathcal{D}_o|\boldsymbol{\Theta}) = \ln \sum_{\mathcal{D}_h} \tilde{p}(\mathcal{D}_o, \mathcal{D}_h|\boldsymbol{\Theta}) \tag{5.5}$$

The quantity ℓ is called the log-likelihood. This alternative way of expressing the ML estimator is mathematically convenient when the dataset $\mathcal{D}_o$ is assumed to be formed by a collection of statistically independent observations. However, the equations for the ML estimator (both alternatives) can be analytically resolved only in some simplified special cases. For example, when the probabilistic model does not assume hidden variables and the observations are assumed to be iid (identically independently distributed) following a Gaussian distribution. For most complex probability models an analytical expression for the ML estimator can not be derived. In those cases a strategy is direct numerical optimization, but this usually implies huge computational costs.

An efficient alternative way for doing this optimization is EM. Since $\tilde{p}(\mathcal{D}_o|\mathbf{\Theta})$ has no dependence on $\mathcal{D}_h$ we can write

$$\begin{aligned}\ell(\mathbf{\Theta};\mathcal{D}_o) &= \ln\tilde{p}(\mathcal{D}_o|\mathbf{\Theta})\cdot 1\\ &= \ln\tilde{p}(\mathcal{D}_o|\mathbf{\Theta})\cdot\sum_{\mathcal{D}_h}\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}')\\ &= \sum_{\mathcal{D}_h}\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}')\cdot\ln\tilde{p}(\mathcal{D}_o|\mathbf{\Theta})\end{aligned}$$

On the other hand, $p(h,o|t) = p(h|o,t)\cdot p(o|t)$ in general. Therefore, $p(o|t) = p(o,h|t)/p(h|o,t)$. Using this in previous equation we get

$$\ell(\mathbf{\Theta};\mathcal{D}_o) = \sum_{\mathcal{D}_h}\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}')\cdot\ln\frac{\tilde{p}(\mathcal{D}_h,\mathcal{D}_o|\mathbf{\Theta})}{\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta})} \tag{5.6}$$

Now let us assume that we have an iterative algorithm that updates the parameters $\mathbf{\Theta}$ in each step. The increase in the log-likelihood in step $n+1$ is

$$\ell(\mathbf{\Theta}^{n+1}) - \ell(\mathbf{\Theta}^n) = \sum_{\mathcal{D}_h}\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}')\cdot\left[\ln\frac{\tilde{p}(\mathcal{D}_h,\mathcal{D}_o|\mathbf{\Theta^{n+1}})}{\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}^{n+1})} - \ln\frac{\tilde{p}(\mathcal{D}_h,\mathcal{D}_o|\mathbf{\Theta}^n)}{\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}^n)}\right]$$

Choosing $\mathbf{\Theta}' = \mathbf{\Theta}^n$ and rearranging terms we get

$$\ell(\mathbf{\Theta}^{n+1}) - \ell(\mathbf{\Theta}^n) = D_{KL} + \sum_{\mathcal{D}_h}\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}^n)\cdot\ln\frac{\tilde{p}(\mathcal{D}_h,\mathcal{D}_o|\mathbf{\Theta^{n+1}})}{\tilde{p}(\mathcal{D}_h,\mathcal{D}_o|\mathbf{\Theta}^n)} \tag{5.7}$$

The term D_{KL} is non-negative since it is the Kullback-Leibler divergence from $\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}^{n+1})$ to $\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}^n)$, which is always ≥ 0

$$D_{KL} = \sum_{\mathcal{D}_h}\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}^n)\cdot\ln\frac{\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}^n)}{\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}^{n+1})}$$

The second term, when maximized with respect to $\mathbf{\Theta}^{n+1}$, is always non-negative since $\mathbf{\Theta}^{n+1} = \mathbf{\Theta}^n$ makes it equal to 0.

In summary, the following iterative algorithm guarantees that the log-likelihood never decreases at each step:

$$\mathbf{\Theta}^{n+1} \leftarrow \arg\max_{\mathbf{\Theta}}\sum_{\mathcal{D}_h}\tilde{p}(\mathcal{D}_h|\mathcal{D}_o,\mathbf{\Theta}^n)\cdot\ln\tilde{p}(\mathcal{D}_h,\mathcal{D}_o|\mathbf{\Theta}) \tag{5.8}$$

This leads to the EM algorithm:

1. **Initialization:** Choose (random) initial parameters $\mathbf{\Theta}^0$. $n \leftarrow 0$
2. **Expectation step:** Calculate $\alpha_{\mathcal{D}_h} = \tilde{p}(\mathcal{D}_h|\mathcal{D}_o, \mathbf{\Theta}^n)$
3. **Maximization step:**

$$\mathbf{\Theta}^{n+1} \leftarrow \arg\max_{\mathbf{\Theta}} \sum_{\mathcal{D}_h} \alpha_{\mathcal{D}_h} \cdot \ln \tilde{p}(\mathcal{D}_h, \mathcal{D}_o|\mathbf{\Theta})$$

4. **Stop criterium:** If the stop criterium is not satisfied go back to 2

In case it can be assumed that the dataset is a collection of N identically independently distributed observations $\{\mathbf{x}_i\}_{i=1}^N$ the EM algorithm can be simplified:

1. **Initialization:** Choose (random) initial parameters $\mathbf{\Theta}^0$. $n \leftarrow 0$
2. **Expectation step:** Calculate $\alpha_{i,h} = \tilde{p}(h|\mathbf{x}_i, \mathbf{\Theta}^n)$
3. **Maximization step:**

$$\mathbf{\Theta}^{n+1} \leftarrow \arg\max_{\mathbf{\Theta}} \sum_{i=1}^{n} \sum_{h} \alpha_{i,h} \cdot \ln \tilde{p}(h, \mathbf{x_i}|\mathbf{\Theta})$$

4. **Stop criterium:** If the stop criterium is not satisfied go back to 2

5.3 Extending EM for MAP parameter estimation

Let us assume that we have a set of observed data $\mathcal{D}_o$. According to the Maximum a Posteriori (MAP) parameter estimation strategy, the optimal values for the parameters $\mathbf{\Theta}$ of the probabilistic model $\{\tilde{p}(\mathcal{D}_o|\mathbf{\Theta}), \tilde{p}(\mathbf{\Theta})\}$ are

$$\hat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}} \mathcal{L}'(\mathbf{\Theta}) \tag{5.9}$$

with $\mathcal{L}'(\mathbf{\Theta})$ being

$$\mathcal{L}'(\mathbf{\Theta}; \mathcal{D}_o) \equiv \tilde{p}(\mathcal{D}_o|\mathbf{\Theta}) \cdot \tilde{p}(\mathbf{\Theta}) \tag{5.10}$$

The hidden (unobserved) variables are taken into account using the law of total probability

$$\mathcal{L}'(\mathbf{\Theta}; \mathcal{D}_o) \equiv \tilde{p}(\mathcal{D}_o|\mathbf{\Theta}) \cdot \tilde{p}(\mathbf{\Theta}) = \left[\sum_{\mathcal{D}_h} \tilde{p}(\mathcal{D}_o, \mathcal{D}_h|\mathbf{\Theta})\right] \cdot \tilde{p}(\mathbf{\Theta}) \tag{5.11}$$

Therefore the MAP estimator is determined by equations 5.9 and 5.11.

Since the logarithm is a monotonically increasing function the MAP estimator equations can be rewritten in an alternative way as

$$\hat{\boldsymbol{\Theta}} = \arg\max_{\boldsymbol{\Theta}} \ell'(\boldsymbol{\Theta}) \tag{5.12}$$

with

$$\ell'(\boldsymbol{\Theta}; \mathcal{D}_o) \equiv \ell(\boldsymbol{\Theta}; \mathcal{D}_o) + \ln \tilde{p}(\boldsymbol{\Theta}) \tag{5.13}$$

Using equation 5.7 we get

$$\ell'(\boldsymbol{\Theta}^{n+1}) - \ell'(\boldsymbol{\Theta}^n) = D_{KL} + \sum_{\mathcal{D}_h} \tilde{p}(\mathcal{D}_h|\mathcal{D}_o, \boldsymbol{\Theta}^n) \cdot \ln \frac{\tilde{p}(\mathcal{D}_h, \mathcal{D}_o|\boldsymbol{\Theta}^{\mathbf{n+1}})}{\tilde{p}(\mathcal{D}_h, \mathcal{D}_o|\boldsymbol{\Theta}^{\mathbf{n}})} + \ln \frac{\tilde{p}(\boldsymbol{\Theta}^{\mathbf{n+1}})}{\tilde{p}(\boldsymbol{\Theta}^{\mathbf{n}})} \tag{5.14}$$

As before, the term D_{KL} is non-negative since it is the Kullback-Leibler divergence from $\tilde{p}(\mathcal{D}_h|\mathcal{D}_o, \boldsymbol{\Theta}^{n+1})$ to $\tilde{p}(\mathcal{D}_h|\mathcal{D}_o, \boldsymbol{\Theta}^n)$. The sum of the second and third terms, when maximized with respect to $\boldsymbol{\Theta}^{n+1}$, is always non-negative since $\boldsymbol{\Theta}^{n+1} = \boldsymbol{\Theta}^n$ makes it equal to 0.

The following iterative algorithm then guarantees that the extended log-likelihood never decreases at each step:

$$\boldsymbol{\Theta}^{n+1} \leftarrow \arg\max_{\boldsymbol{\Theta}} \left[\ln \tilde{p}(\boldsymbol{\Theta}) + \sum_{\mathcal{D}_h} \tilde{p}(\mathcal{D}_h|\mathcal{D}_o, \boldsymbol{\Theta}^n) \cdot \ln \tilde{p}(\mathcal{D}_h, \mathcal{D}_o|\boldsymbol{\Theta}) \right] \tag{5.15}$$

This leads to the following extended EM algorithm that performs MAP parameter estimation:

1. **Initialization:** Choose (random) initial parameters $\boldsymbol{\Theta}^0$. $n \leftarrow 0$

2. **Expectation step:** Calculate $\alpha_{\mathcal{D}_h} = \tilde{p}(\mathcal{D}_h|\mathcal{D}_o, \boldsymbol{\Theta}^n)$

3. **Maximization step:**

$$\boldsymbol{\Theta}^{n+1} \leftarrow \arg\max_{\boldsymbol{\Theta}} \left[\ln \tilde{p}(\boldsymbol{\Theta}) + \sum_{\mathcal{D}_h} \alpha_{\mathcal{D}_h} \cdot \ln \tilde{p}(\mathcal{D}_h, \mathcal{D}_o|\boldsymbol{\Theta}) \right] \tag{5.16}$$

4. **Stop criterium:** If the stop criterium is not satisfied go back to 2

5.4 Multiclustering

Let us assume that our dataset is composed by M subsets

$$\mathcal{D}_o = \left\{\mathcal{D}_o^1, ..., \mathcal{D}_o^M\right\} \tag{5.17}$$

with subset j composed by N^j observations

$$\mathcal{D}_o^j = \{\mathbf{x}_{j,i}\}_{i=1}^{N^j} \tag{5.18}$$

Our statistical model assumes that the observations in a given subset are iid following a Gaussian mixture of K components; the priors of the components in subset j are $\{\pi_{j,c}\}_{c=1}^K$; the centers of the components in subset j are $\{\mathbf{m}_{j,c}\}_{c=1}^K$ and their covariance matrices are $\{\mathbf{C}_c\}_{c=1}^K$. Therefore the covariance matrix of component c is assumed to be the same across all the subsets. The probability distribution in the whole dataset is then

$$\tilde{p}(\mathcal{D}_h, \mathcal{D}_o|\mathbf{\Theta}) = \prod_{j=1}^{M}\prod_{i=1}^{N^j} \tilde{p}\left(h_i, \mathbf{x}_i|\left\{\pi_{j,c}, \mathbf{m}_{j,c}, \mathbf{C}_c\right\}_{c=1}^K\right) \tag{5.19}$$

Now we can rewrite the last term of 5.16 as

$$\begin{aligned}\sum_{\mathcal{D}_h} \alpha_{\mathcal{D}_h} \cdot \ln \tilde{p}(\mathcal{D}_h, \mathcal{D}_o|\mathbf{\Theta}) &= \sum_{\mathcal{D}_h} \alpha_{\mathcal{D}_h} \cdot \sum_{j=1}^{M}\sum_{i=1}^{N^j} \ln \tilde{p}\left(h_i, \mathbf{x}_i|\left\{\pi_{j,c}, \mathbf{m}_{j,c}, \mathbf{C}_c\right\}_{c=1}^K\right) \\ &= \sum_{j=1}^{M}\sum_{i=1}^{N^j}\sum_{h} \alpha_{i,j,c} \cdot \ln \tilde{p}\left(c, \mathbf{x}_i|\left\{\pi_{j,c}, \mathbf{m}_{j,c}, \mathbf{C}_c\right\}_{c=1}^K\right)\end{aligned}$$

with $\alpha_{i,j,c} = \tilde{p}\left(c\,|\,\mathbf{x}_{j,i}, \left\{\pi_{j,c}^n, \mathbf{m}_{j,c}^n, \mathbf{C}_c^n\right\}_{c=1}^K\right)$

Now we will discuss the meaning of the log-prior $\ln \tilde{p}(\mathbf{\Theta})$ in equation 5.15 and propose a prior for our multiclustering problem. A non uniform prior reflects that some parameter configurations are in principle preferred to others. In our case we would like that the component means $\mathbf{m}_{j,c}$ do not vary "too much" across the subsets, that is, $\mathbf{m}_{j,c} \sim \mathbf{m}_{j',c}$. We represent this preferrence as

$$\mathbf{m}_{j,c} = \mathbf{m}_c + \mathbf{e}_{j,c} \tag{5.20}$$

and

$$\ln \tilde{p}(\mathbf{\Theta}) = -\gamma \sum_{j,c} w_j \, D_{KL}\left(\mathcal{G}_c||\mathcal{G}_{j,c}\right) \tag{5.21}$$

where $\mathcal{G}_{j,c}$ is the probability distribution of component c in subset j (which has mean $\mathbf{m}_{j,c}$ and covariance matrix $\mathbf{C}_c$) and $\mathcal{G}_c$ represents a consensum of the different

probability distributions of component c across the subsets (mean $\mathbf{m}_c$ and covariance matrix $\mathbf{C}_c$). The factor w_j is a parameter that weights the influence of each subset in the log-prior. It is straightforward to show that the log-prior is

$$\ln \tilde{p}(\mathbf{\Theta}) = -\frac{\gamma}{2} \sum_{j,c} w_j \left\{ [\mathbf{m}_{j,c} - \mathbf{m}_c]^T \mathbf{C}_c^{-1} [\mathbf{m}_{j,c} - \mathbf{m}_c] + d + d \ln 2\pi + |\mathbf{C}_c| \right\} \quad (5.22)$$

5.4.1 Derivation of the specific equations

A closed-form expression for the maximization of 5.16 can be derived in some cases. Since we are assuming that the components are characterized by Gaussian distributions we get that the functional to maximize is

$$\begin{aligned} f &= -\frac{\gamma}{2} \sum_{j,c} w_j \left\{ [\mathbf{m}_{j,c} - \mathbf{m}_c]^T \mathbf{C}_c^{-1} [\mathbf{m}_{j,c} - \mathbf{m}_c] + d + d \ln 2\pi + \ln |\mathbf{C}_c| \right\} \\ &+ \sum_{i,j,c} \alpha_{i,j,c} \cdot \left[\ln \pi_{j,c} - \frac{d}{2} \ln 2\pi + \frac{1}{2} \ln |\mathbf{C}_c^{-1}| - \frac{1}{2} [\mathbf{x}_{j,i} - \mathbf{m}_{j,c}]^T \mathbf{C}_c^{-1} [\mathbf{x}_{j,i} - \mathbf{m}_{j,c}] \right] \\ &+ \sum_j \lambda_j \left(1 - \sum_c \pi_{j,c} \right) \end{aligned}$$

We introduced the last term in order to take into account the constraint $\sum_c \pi_{j,c} = 1$. In the appendix we derive the following expressions for the parameter values that maximize the functional:

$$\begin{aligned} \pi_{j,c}^{n+1} &= \frac{1}{N_j} \sum_i \alpha_{i,j,c} \\ \mathbf{m}_c^{n+1} &= \left[\sum_j \frac{\gamma\, w_j \sum_i \alpha_{i,j,c}}{\gamma\, w_j + \sum_i \alpha_{i,j,c}} \right]^{-1} \sum_j \frac{\gamma\, w_j}{\gamma\, w_j + \sum_i \alpha_{i,j,c}} \sum_i \alpha_{i,j,c}\, \mathbf{x}_{j,i} \\ \mathbf{m}_{j,c}^{n+1} &= \frac{1}{\gamma\, w_j + \sum_i \alpha_{i,j,c}} \left[\gamma\, w_j\, \mathbf{m}_c^{n+1} + \sum_i \alpha_{i,j,c}\, \mathbf{x}_{j,i} \right] \\ \mathbf{C}_c^{n+1} &= \mathbf{A}_c + \mathbf{B}_c \end{aligned}$$

with

$$\begin{aligned} \mathbf{A}_c &= \frac{1}{\gamma \sum_j w_j + \sum_{i,j} \alpha_{i,j,c}} \sum_j \gamma\, w_j \left[\mathbf{m}_{j,c}^{n+1} - \mathbf{m}_c^{n+1} \right] \left[\mathbf{m}_{j,c}^{n+1} - \mathbf{m}_c^{n+1} \right]^T \\ \mathbf{B}_c &= \frac{1}{\gamma \sum_j w_j + \sum_{i,j} \alpha_{i,j,c}} \sum_{i,j} \alpha_{i,j,c} \left[\mathbf{x}_{j,i} - \mathbf{m}_{j,c}^{n+1} \right] \left[\mathbf{x}_{j,i} - \mathbf{m}_{j,c}^{n+1} \right]^T \end{aligned}$$

5.4.2 Special cases

In order to check our equations we will consider two special cases:

Special case 1: Only one subset ($M = 1$)

In this case our equations simplify to:

$$\begin{aligned}
\pi_c^{n+1} &= \frac{1}{N} \sum_i \alpha_{i,c} \\
\mathbf{m}_c^{n+1} &= \frac{1}{\sum_i \alpha_{i,c}} \sum_i \alpha_{i,c}\, \mathbf{x}_i \\
\mathbf{C}_c^{n+1} &= \frac{1}{\gamma\, w + \sum_i \alpha_{i,c}} \sum_i \alpha_{i,c} \left[\mathbf{x}_i - \mathbf{m}_c^{n+1}\right] \left[\mathbf{x}_i - \mathbf{m}_c^{n+1}\right]^T
\end{aligned}$$

For $\gamma = 0$ these equations coincide with the standard EM equations for a mixture of Gaussians.

Special case 2: $\gamma \to \infty$

In this case $\mathbf{C}_c^{n+1} \to \mathbf{0}$, so $\alpha_{i,j,c} \to 0, 1$. The algorithm then gets:

$$\begin{aligned}
\mathbf{m}_c^{n+1} &= \frac{1}{\sum_j N_{j,c}} \sum_j N_{j,c}\, \mu_{j,c} \\
\mathbf{m}_{j,c}^{n+1} &= \mathbf{m}_c^{n+1}
\end{aligned}$$

which is equivalent to k-means applied to a single set formed by the union of all the subsets.

Special case 3: covariance matrices $\to 0$

In this case the $\alpha_{i,j,c} \to 0, 1$. The algorithm gets:

$$\begin{aligned}
\mathbf{m}_c^{n+1} &= \left[\sum_j \frac{\gamma\, w_j N_{j,c}}{\gamma\, w_j + N_{j,c}}\right]^{-1} \sum_j \frac{\gamma\, w_j N_{j,c}}{\gamma\, w_j + N_{j,c}}\, \mu_{j,c} \\
\mathbf{m}_{j,c}^{n+1} &= \frac{1}{\gamma\, w_j + N_{j,c}} \left[\gamma\, w_j\, \mathbf{m}_c^{n+1} + N_{j,c}\, \mu_{j,c}\right]
\end{aligned}$$

5.5 Results

We investigated the performance of multiclustering algorithm in the analysis of synthetic data and MRI data. This section shows the results obtained in each case. In addition, the results of our algorithm are compared with the results obtained by applying two conventional clustering algorithms to analyze the same data. All the images where provided as a single dataset to the standard algorithms in order to find the common patterns.

5.5.1 Synthetic examples

Two experiments with synthetic data were performed using different 2D-Gaussian distribution setups. First, random samples where generated from five different 2D-Gaussian distributions. The samples from each distribution where grouped into different datasets, one per distribution, so each dataset represents a synthetic cluster (Figure 5.2). On the second experiment random samples where generated from three different 2D-Gaussian, but in this case datatasets contain samples from each of the distributions but with different priors, which is the probability of getting samples from each specific cluster (Figure 5.3). This is the expected case for real scenarios, where some subjects will just present a few of the natural clusters.

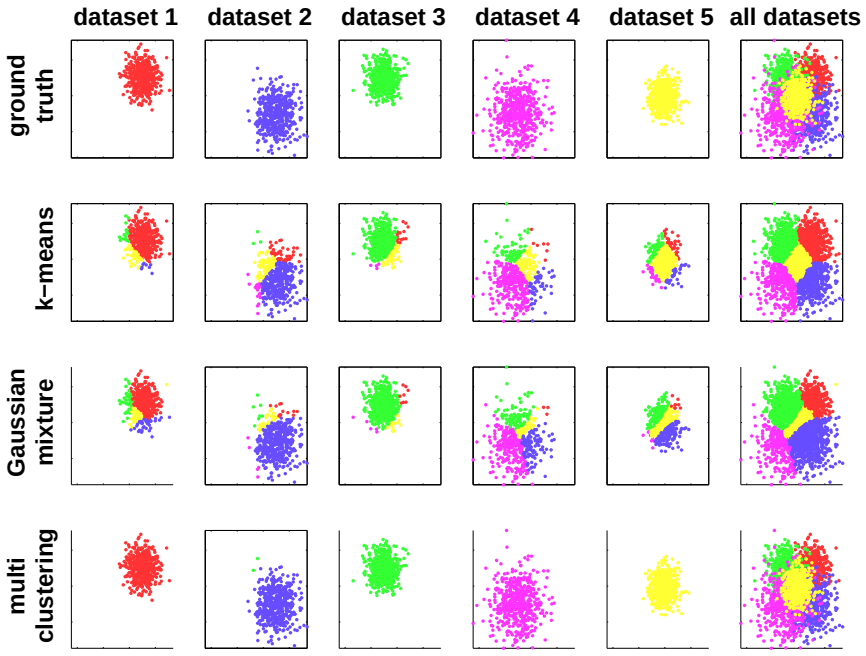

Figure 5.2: Synthetic example 1 with different clusters in each dataset. First row presents the natural structure of the simulated data and each other row shows the pixels that are assigned to each of the five clusters that algorithms are configured to find.

For the first experiment, Figure 5.2 gathers the results using different clustering algorithms; k-means, Gaussian mixture with standard EM (GMM-EM), and multiclustering. When compared with the ground truth, located at the first row of the figure, it is observed that our algorithm is able to successfully detect the natural structure of the data, internally determining that each dataset exposes a different cluster. On the other hand, k-means and GMM-EM algorithms, which are not aware that there are different sources of data, can only find clusters in the complete mix of datasets, which is why multiple cluster are incorrectly found in each source dataset.

The results of the second experiment, is shown in figure 5.3, where the same algorithms are applied. Again, our algorithm overcomes the standard clustering methods. Now different clusters are found in each of the source datasets, and the priors are close to the simulated values in the sense that the proportion of samples assigned to each cluster is similar to the one in the ground truth. On the other hand, the other methods are incorrectly assigning the observations, specially because the proportion of the samples assigned to each cluster does not vary much from each source dataset to the other.

Figure 5.4 shows the convergence behavior of multiclustering applied to this dataset.

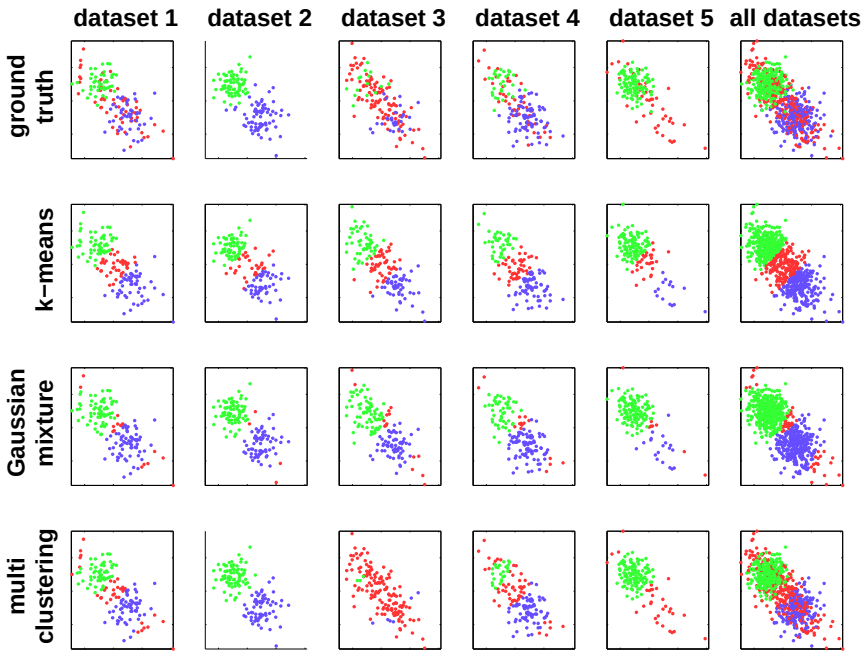

Figure 5.3: Synthetic example 2 simulated with three different clusters, which are not necessarily in all datasets. First row presents the natural structure of the simulated data and each other row shows the pixels that are assigned to each of the three clusters that algorithms are configured to find.

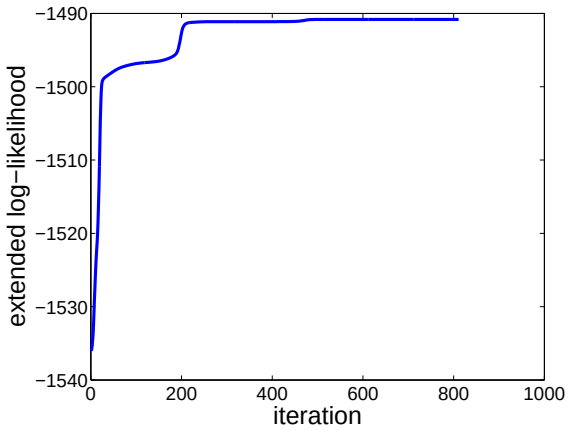

Figure 5.4: Multiclustering algorithm convergence. The extended log-likelihood function, that is internally maximized by the multiclustering algorithm, is constantly increased until final convergence is reached.

5.5.2 MRI data

Multiclustering algorithm was also applied to MRI data obtained from three studies of glioma development in animal models as a prove of concept. A detailed description of the acquisition and pre-processing of the analyzed data can be found in the section 3.2. To show our algorithm performance in a simple way, and based on the expert knowledge of the structure of the analyzed data, the algorithms were configured to find two cluster.

First study corresponds to twenty glioma rat models, ten of high-grade and ten of low-grade (Tumor_1 dataset). From each subject two MR images, of types MT and T_{2w}, where acquired simultaneously. Hence each pixel has two dimensions, one per image. For the multiclustering case, the source of each pixel was also provided. Table 5.1 and Figure 5.5 show the obtained cluster priors using k-means, GMM-EM, and multiclustering algorithm in the Tumor_1 dataset. This indicates the proportion of pixels that are assigned to each cluster for each subject.

The goal of the analysis is to find if there is a pattern, in the form of a cluster, that is predominantly related to the grade of the tumor which is known but not provided to the clustering algorithms. In this case, k-means and multiclustering create clusters that are highly correlated to the tumor types, for instance, multiclustering assigns higher priory values to C_2 on low-graded gliomas images and higher priory values of C_1 on high-graded gliomas images. In this case GMM-EM does not present any recognizable pattern.

Second study corresponds to twelve high-grade glioma rat model divided in two subgroups of six male and six female (Tumor_3 dataset). From each subject fourteen DWI where acquired, with two different b-values per seven directions. Now each pixel has fourteen dimensions, one per image. Again, algorithms where configured

to find two clusters.

Table 5.2 and Figure 5.6 present priors found with the algorithms. In this case the multiclustering priors are completely related to the gender, where all female subjects present more pixels in C_1 and all male subjects locate the majority of the pixels at C_2. k-means also finds high correlation between gender and one of the created clusters, with two exceptions, and GMM-EM provides the poorer results again.

Finally, the third study corresponds to ten mouse models of glioblastoma undergoing an antiangiogenic therapy, six of them responded favorably to treatment and the other four did not (Tumor_4 dataset). In this case each subject provides nine DWI, with three different b-values per three orthogonal directions. Each pixel has nine dimensions, one per image. Again, algorithms where configured to find two clusters.

Table 5.3 and Figure 5.7 present priors found with the algorithms. Again multiclustering algorithm presents the best performance, because the priors are completely related to the type of therapy response, higher C_1 priors are presented in "response" cases and all higher C_2 priors are presented on "non response" cases. k-means and GMM-EM also finds some correlation between response type and one of the created clusters.

rat	grade	k-means		gmm		multicl.	
		C_1	C_2	C_1	C_2	C_1	C_2
1		0.0	1.0	0.0	1.0	0.0	1.0
2		0.0	1.0	0.0	1.0	0.0	1.0
3		0.4	0.6	1.0	0.0	0.6	0.4
4		0.0	1.0	0.0	1.0	0.0	1.0
5	**low**	0.0	1.0	0.0	1.0	0.1	0.9
6		0.0	1.0	0.0	1.0	0.0	1.0
7		0.0	1.0	0.4	0.6	0.1	0.9
8		0.1	0.9	0.0	1.0	0.1	0.9
9		0.0	1.0	0.1	0.9	0.0	1.0
10		0.4	0.6	1.0	0.0	0.3	0.7
11		0.4	0.6	0.0	1.0	0.6	0.4
12		0.6	0.4	0.9	0.1	0.8	0.2
13		0.7	0.3	1.0	0.0	0.5	0.5
14		0.5	0.5	0.2	0.8	0.7	0.3
15	**high**	0.8	0.2	0.8	0.2	0.4	0.6
16		0.3	0.7	1.0	0.0	0.6	0.4
17		0.7	0.3	0.0	1.0	0.1	0.9
18		0.8	0.2	1.0	0.0	0.4	0.6
19		0.9	0.1	1.0	0.0	0.5	0.5
20		0.9	0.1	1.0	0.0	0.9	0.1

Table 5.1: Clusters obtained for the database of glioma grades. $K = 2$. For each subject, identified by the first column, the table presents the pixel proportion that is related to each of the found clusters by the algorithms indicated in the first row.

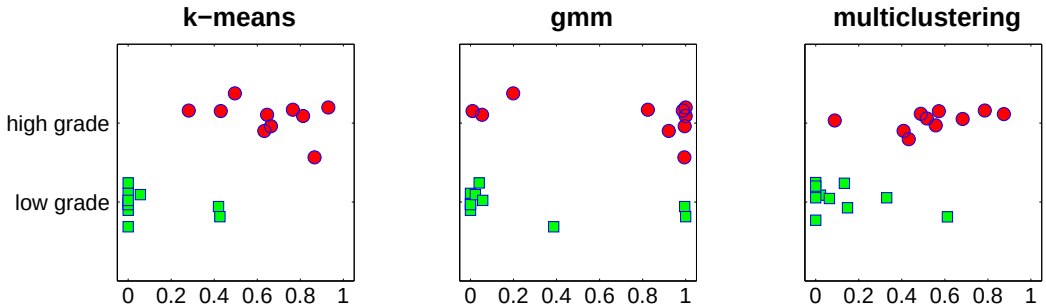

Figure 5.5: Presence of cluster 1 in the different subjects of the glioma grade database. $K = 2$. It is appreciated that k-means and multiclustering associate more pixels to cluster C_1 when the subject presents a high degree tumor.

rat	gender	k-means		gmm		multicl.	
		C_1	C_2	C_1	C_2	C_1	C_2
1		0.0	1.0	1.0	0.0	0.7	0.3
2		0.1	0.9	1.0	0.0	1.0	0.0
3	**female**	0.0	1.0	0.0	1.0	0.8	0.2
4		0.2	0.8	0.0	1.0	0.8	0.2
5		0.3	0.7	0.0	1.0	0.8	0.2
6		1.0	0.0	0.0	1.0	1.0	0.0
7		0.0	1.0	0.0	1.0	0.0	1.0
8		0.0	1.0	0.0	1.0	0.0	1.0
9	**male**	0.0	1.0	0.0	1.0	0.0	1.0
10		0.0	1.0	0.0	1.0	0.1	0.9
11		0.0	1.0	0.0	1.0	0.0	1.0
12		0.0	1.0	0.0	1.0	0.0	1.0

Table 5.2: Clusters obtained for the database of gliomas across genders. $K = 2$**.** For each subject, identified by the first column, the table presents the pixel proportion that is related to each of the found clusters by the algorithms indicated in the first row.

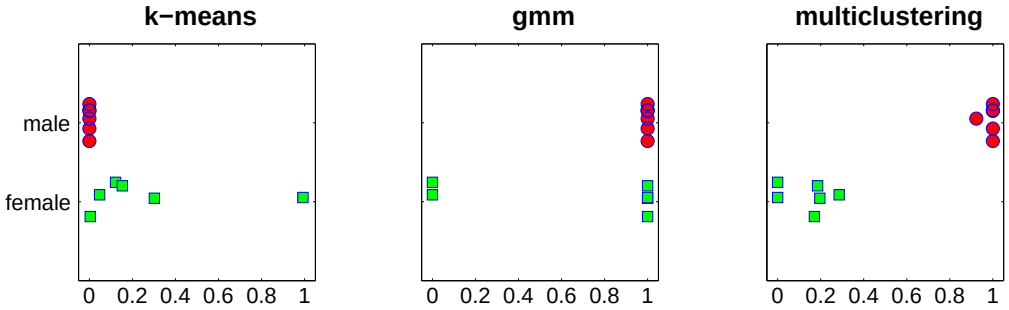

Figure 5.6: Presence of cluster 1 **in the different subjects of the gliomas across genders database.** $K = 2$**.** It is appreciated that multiclustering associate almost all the pixels to cluster C_1 for male subjects.

rat	evolution	k-means		gmm		multicl.	
		C_1	C_2	C_1	C_2	C_1	C_2
1		1.0	0.0	0.0	1.0	0.1	0.9
2		0.8	0.2	0.0	1.0	0.1	0.9
3	**no response**	1.0	0.0	0.4	0.6	0.3	0.7
6		1.0	0.0	0.0	1.0	0.0	1.0
4		1.0	0.0	0.4	0.6	0.4	0.6
5		0.7	0.3	0.3	0.7	0.5	0.5
7		1.0	0.0	0.0	1.0	0.6	0.4
8		0.9	0.1	0.0	1.0	0.4	0.6
9	**response**	0.6	0.4	0.4	0.6	0.6	0.4
10		0.0	1.0	1.0	0.0	1.0	0.0
11		0.3	0.7	1.0	0.0	1.0	0.0
12		0.4	0.6	0.5	0.5	0.6	0.4

Table 5.3: Clusters obtained for the database of response to anti-angiogenic treatment. $K = 2$. For each subject, identified by the first column, the table presents the pixel proportion that is related to each of the found clusters by the algorithms indicated in the first row.

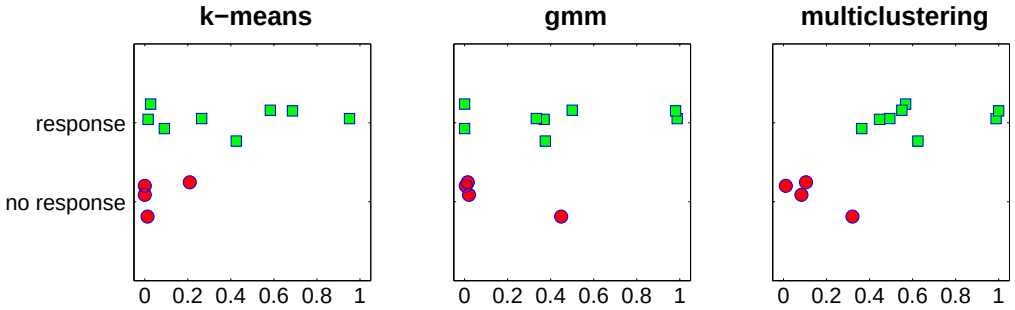

Figure 5.7: Presence of cluster 1 in the different subjects of the anti-angiogenic treatment database. $K = 2$. It is appreciated that multiclustering associate more pixels to cluster C_1 for response subjects.

5.5.3 Pattern visualization on MRI

The clustering methodology allows the visualization of the localization of the different clusters within the investigated images, tracing back the original positions of the pixels, once assigned by the algorithm. Importantly, visual inspection of the cluster localization could be of help to oncologists by highlighting diagnostic information of the tumors, such as areas of proliferation and growth or regression.

To illustrate this idea, the multiclustering algorithm was applied to brain tumors in human patients. Specifically, MRI data were analyzed from 35 patients with gliomas, 17 low grade and 18 high grade, studied through images of types T_{1w} FLAIR post-gadolinium and T_{1w} pre-gadolinium [FNLCR, 2014]. A more detailed description of the acquisition and pre-processing of these data may be found in section 3.2. For this analysis we configured the algorithm to find and locate three clusters in the database, based on expert anatomical assessment of tumor morphology.

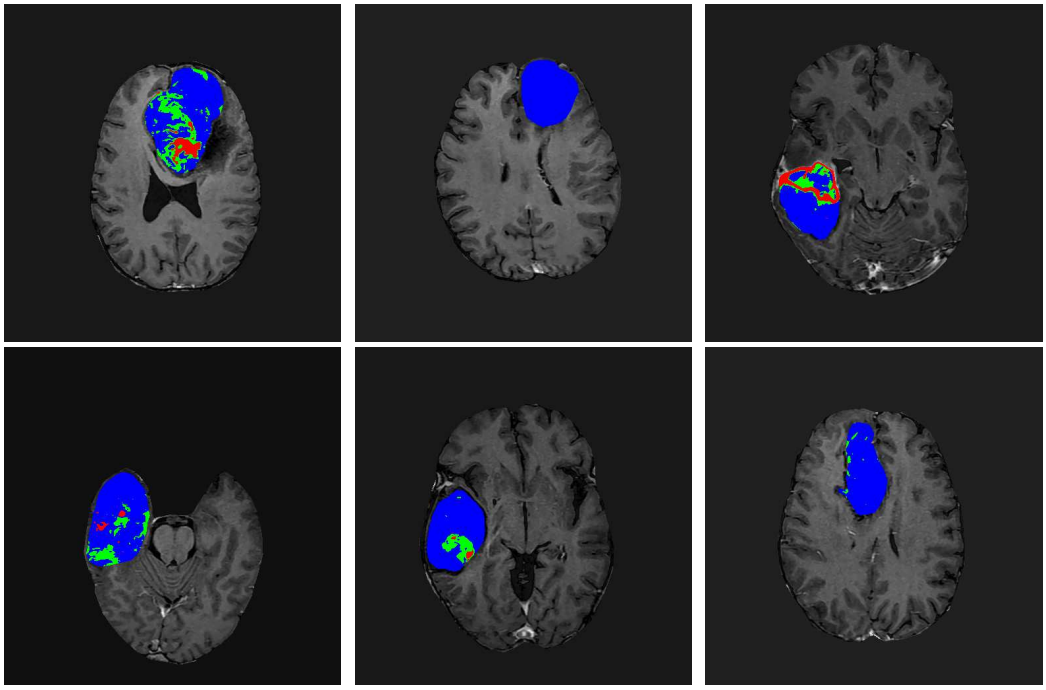

Figure 5.8: Visualization of clusters on MRI scans from representative patients with low grade gliomas. Clusters are superimposed over a T_{1w} FLAIR post-gadolinium image. Blue: cluster associates with low or absent proliferation. Green: cluster associates with intermediate proliferation. Red: cluster associates with high proliferation.

Figures 5.8 and 5.9 show the pixels of each cluster on a representative selection of the original images of the T_{1w} FLAIR post-gadolinium dimension. The pixels assigned to each cluster are represented by three different colors. The blue cluster can be associated to areas of reduced growth and proliferation, the green areas to regions of intermediate growth or edema, and the red areas to fast growing tumor. It

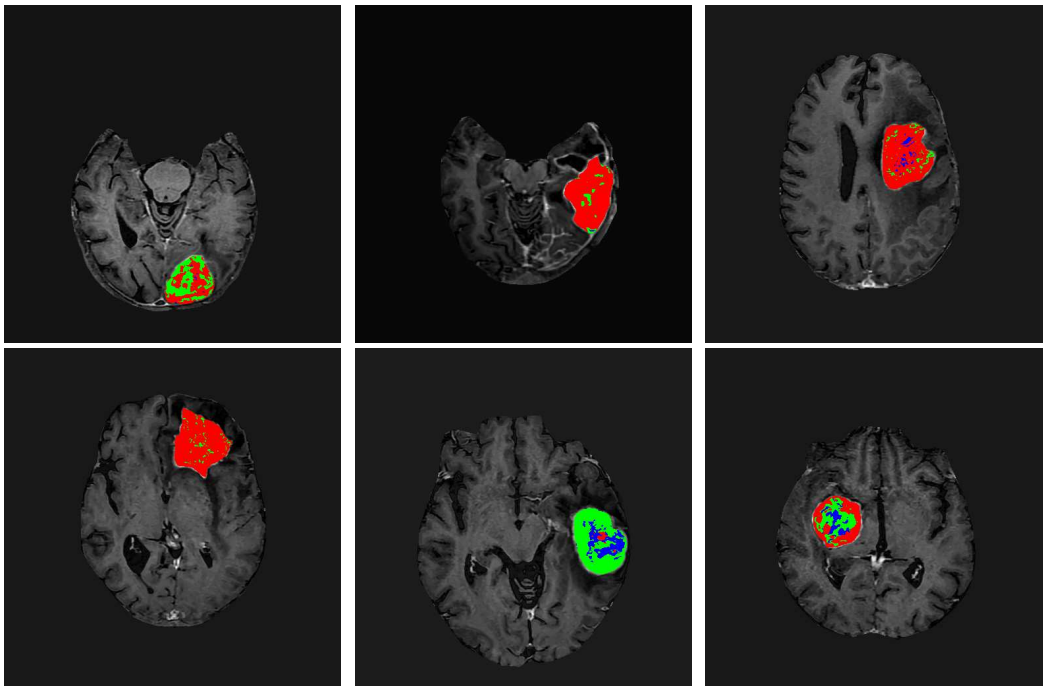

Figure 5.9: Visualization of clusters on MRI scans from representative patients with high grade gliomas. Clusters are superimposed over a T_{1w} FLAIR post-gadolinium image. Blue: cluster associates with low or absent proliferation. Green: cluster associates with intermediate proliferation. Red: cluster associates with high proliferation.

is to note that in some cases, even low grade tumors show small zones of intermediate (green) and fast (red) growth. In high grade glioma tumors, most tumors depict large green and red areas, indicating they are dominated by intense proliferation zones, while blue areas indicating reduced or absent growth are very small, or even inexistent.

5.6 Discussion

In this chapter, we proposed a novel clustering algorithm to find common patterns in small databases, based in the joint analysis of all available data, but being also able to identify individual contributions. Compared with standard algorithms, our algorithm presented better results both, in synthetic data and in data from MRI studies.

The two clustering algorithms chosen as benchmarks, have been previously used to analyze MRI databases with valuable results [Nguyen et al., 2015, Fontana et al., 2016, Nekooeimehr et al., 2016, Triadyaksa et al., 2016, Donoso et al., 2010, Simon et al., 2012, Oikonomou and Blekas, 2013, Lu et al., 2013]. The fact that our multiclustering algorithm outperforms earlier approaches provides a measure of

confidence on its future developments.

For simplicity and based on expert knowledge, we configured the algorithms to find two clusters. However, the proposed algorithm is not limited to a specific number of clusters, but it is still necessary to implement an automatic method to select the optimum number of clusters that best suits the data. Finding the optimal number of clusters, may help additionally to find specific patterns that explain specific characteristics of the tumors that make them resistant to certain therapies.

This algorithm has been used to extract relevant information in experimental studies with few subjects. Nevertheless, the potential applications of this multiclustering algorithm, are not restricted, or limited to, problems with small numbers of subjects. In fact, the proposed algorithm is a general purpose method, which can be applied to larger databases, and to other problems not related to image analysis.

Concerning the results obtained, we were able to find useful patterns to distinguish automatically between high grade and low grade gliomas. These may become very useful to support important clinical decision with enormous therapeutic and prognostic consequences.

Similarly, we found patterns related to the sexual dimorphism of these tumors, which may probably help to understand why male tumors are more morbid and malign than female tumors.

It is noteworthy that the algorithm proposed was able to identify patterns specifically related to "responders" and to "not responders", even only after two sessions of anti-angiogenic therapy. Using these patterns to study "responders" and "non-responders" to anti-angiogenic therapies entails crucial consequences, since administering the anti-angiogenic therapy to a non-responder represents not only a waste of economic resources, but also a waste of time to apply a more effective therapy. These decisions are normally based on previous experience by the medical team, but the development of diagnostic support environments would help enormously to strengthen the medical decision, or even support decisions in distant places where medical advice is unavailable. This opens the path to analyze time-dependent changes and temporal evolution of tumoral, and non tumoral diseases. The extension to neurodegenerative pathologies as Alzheimer disease, Parkinson Disease, Multiple Sclerosis or Lateral Amyotrophic Sclerosis, and their progression with time, constitute a natural extension of the present work.

In summary, the results obtained with our multiclustering algorithm provide a sound basis to foster further applications in image analysis of brain tumor and other pathologies.

5.7 Appendix

The functional to maximize is

$$\begin{aligned} f &= -\frac{\gamma}{2}\sum_{j,c} w_j \left\{ [\mathbf{m}_{j,c} - \mathbf{m}_c]^T \mathbf{C}_c^{-1} [\mathbf{m}_{j,c} - \mathbf{m}_c] + d + d\ln 2\pi + \ln|\mathbf{C}_c| \right\} \\ &+ \sum_{i,j,c} \alpha_{i,j,c} \cdot \left[\ln \pi_{j,c} - \frac{d}{2}\ln 2\pi + \frac{1}{2}\ln|\mathbf{C}_c^{-1}| - \frac{1}{2}[\mathbf{x}_{j,i} - \mathbf{m}_{j,c}]^T \mathbf{C}_c^{-1} [\mathbf{x}_{j,i} - \mathbf{m}_{j,c}] \right] \\ &+ \sum_j \lambda_j \left(1 - \sum_c \pi_{j,c} \right) \end{aligned}$$

Derivative wrt $\pi_{j,c}$

In order to take into account the constraint $\sum_c \pi_{j,c} = 1$ we have to add to the functional f the Lagrangian constraint $\sum_j \lambda_j (1 - \sum_c \pi_{j,c})$. The derivative of the expanded functional wrt $\pi_{j,c}$ is:

$$\sum_i \alpha_{i,j,c} \frac{1}{\pi_{j,c}} - \lambda_j$$

Equating this to zero and rearranging terms gives us:

$$\pi^*_{j,c} = \frac{1}{\lambda^*_j} \sum_i \alpha_{i,j,c}$$

Summing in c in both left and right terms we get:

$$1 = \frac{1}{\lambda^*_j} \sum_{i,c} \alpha_{i,j,c} = \frac{1}{\lambda^*_j} N_j$$

Therefore, $\lambda^*_j = N_j$ and

$$\pi^*_{j,c} = \frac{1}{N_j} \sum_i \alpha_{i,j,c}$$

Derivative wrt $\mathbf{m}_c$

The derivative of the functional f wrt $\mathbf{m}_c$ is

$$\frac{\partial}{\partial \mathbf{m}_c} f = \gamma \sum_c w_j \mathbf{C}_c^{-1} [\mathbf{m}_{j,c} - \mathbf{m}_c]$$

Equating to zero and simplifying we get

$$\mathbf{m}_c^* = \frac{1}{\sum_j w_j} \sum_j w_j \cdot \mathbf{m}_{j,c}^*$$

Derivative wrt $\mathbf{m}_{j,c}$

The derivative of the functional f wrt $\mathbf{m}_{j,c}$ is

$$\frac{\partial}{\partial \mathbf{m}_{j,c}} f = \gamma\, w_j\, \mathbf{C}_c^{-1} \left[\mathbf{m}_{j,c} - \mathbf{m}_c\right] + \sum_i \alpha_{i,j,c} \mathbf{C}_c^{-1} \left[\mathbf{m}_{j,c} - \mathbf{x}_{j,i}\right]$$

Equating to zero and simplifying we get:

$$\mathbf{m}_{j,c}^* = \frac{1}{\gamma\, w_j + \sum_i \alpha_{i,j,c}} \left[\gamma\, w_j\, \mathbf{m}_c^* + \sum_i \alpha_{i,j,c}\, \mathbf{x}_{j,i}\right]$$

Now,

$$\mathbf{m}_c^* = \frac{1}{\sum_j w_j} \sum_j \frac{w_j}{\gamma\, w_j + \sum_i \alpha_{i,j,c}} \left[\gamma\, w_j\, \mathbf{m}_c^* + \sum_i \alpha_{i,j,c}\, \mathbf{x}_{j,i}\right]$$

$$\mathbf{m}_c^* \sum_j w_j = \mathbf{m}_c^* \sum_j \frac{\gamma\, w_j^2}{\gamma\, w_j + \sum_i \alpha_{i,j,c}} + \sum_j \frac{w_j}{\gamma\, w_j + \sum_i \alpha_{i,j,c}} \sum_i \alpha_{i,j,c}\, \mathbf{x}_{j,i}$$

$$\mathbf{m}_c^* \sum_j \frac{w_j \sum_i \alpha_{i,j,c}}{\gamma\, w_j + \sum_i \alpha_{i,j,c}} = \sum_j \frac{w_j}{\gamma\, w_j + \sum_i \alpha_{i,j,c}} \sum_i \alpha_{i,j,c}\, \mathbf{x}_{j,i}$$

That is

$$\mathbf{m}_c^* = \left[\sum_j \frac{w_j \sum_i \alpha_{i,j,c}}{\gamma\, w_j + \sum_i \alpha_{i,j,c}}\right]^{-1} \sum_j \frac{w_j}{\gamma\, w_j + \sum_i \alpha_{i,j,c}} \sum_i \alpha_{i,j,c}\, \mathbf{x}_{j,i}$$

Derivative wrt $\mathbf{C}_c^{-1}$

The derivative of the functional f wrt $\mathbf{C}_c^{-1}$ is

$$\begin{aligned}\frac{\partial}{\partial \mathbf{C}_c^{-1}} f = & - \frac{\gamma}{2} \sum_j w_j \left\{ [\mathbf{m}_{j,c} - \mathbf{m}_c] [\mathbf{m}_{j,c} - \mathbf{m}_c]^T - \mathbf{C}_c^{-1} \right\} \\ & + \frac{1}{2} \sum_{i,j} \alpha_{i,j,c} \left[\mathbf{C}_c - [\mathbf{x}_{j,i} - \mathbf{m}_{j,c}] [\mathbf{x}_{j,i} - \mathbf{m}_{j,c}]^T \right]\end{aligned}$$

Equating to zero and simplifying we get

$$\begin{aligned}\mathbf{C}_c^* = & \frac{1}{\gamma \sum_j w_j + \sum_{i,j} \alpha_{i,j,c}} \sum_j \gamma \, w_j \left[\mathbf{m}_{j,c}^* - \mathbf{m}_c^* \right] \left[\mathbf{m}_{j,c}^* - \mathbf{m}_c^* \right]^T \\ + & \frac{1}{\gamma \sum_j w_j + \sum_{i,j} \alpha_{i,j,c}} \sum_{i,j} \alpha_{i,j,c} \left[\mathbf{x}_{j,i} - \mathbf{m}_{j,c}^* \right] \left[\mathbf{x}_{j,i} - \mathbf{m}_{j,c}^* \right]^T\end{aligned}$$

In summary, we got that the parameter values that optimize the functional f are

$$\begin{aligned}\pi_{j,c}^{n+1} = & \frac{1}{N_j} \sum_i \alpha_{i,j,c} \\ \mathbf{m}_c^{n+1} = & \left[\sum_j \frac{\gamma \, w_j \sum_i \alpha_{i,j,c}}{\gamma \, w_j + \sum_i \alpha_{i,j,c}} \right]^{-1} \sum_j \frac{\gamma \, w_j}{\gamma \, w_j + \sum_i \alpha_{i,j,c}} \sum_i \alpha_{i,j,c} \, \mathbf{x}_{j,i} \\ \mathbf{m}_{j,c}^{n+1} = & \frac{1}{\gamma \, w_j + \sum_i \alpha_{i,j,c}} \left[\gamma \, w_j \, \mathbf{m}_c^{n+1} + \sum_i \alpha_{i,j,c} \, \mathbf{x}_{j,i} \right] \\ \mathbf{C}_c^{n+1} = & \mathbf{A}_c + \mathbf{B}_c\end{aligned}$$

with

$$\begin{aligned}\mathbf{A}_c = & \frac{1}{\gamma \sum_j w_j + \sum_{i,j} \alpha_{i,j,c}} \sum_j \gamma \, w_j \left[\mathbf{m}_{j,c}^{n+1} - \mathbf{m}_c^{n+1} \right] \left[\mathbf{m}_{j,c}^{n+1} - \mathbf{m}_c^{n+1} \right]^T \\ \mathbf{B}_c = & \frac{1}{\gamma \sum_j w_j + \sum_{i,j} \alpha_{i,j,c}} \sum_{i,j} \alpha_{i,j,c} \left[\mathbf{x}_{j,i} - \mathbf{m}_{j,c}^{n+1} \right] \left[\mathbf{x}_{j,i} - \mathbf{m}_{j,c}^{n+1} \right]^T\end{aligned}$$

$$\begin{aligned}\mathbf{m}_{j,c}^{n+1} - \mathbf{m}_c^{n+1} &= \frac{1}{\gamma \, w_j + \sum_i \alpha_{i,j,c}} \left[\gamma \, w_j \, \mathbf{m}_c^{n+1} + \sum_i \alpha_{i,j,c} \, \mathbf{x}_{j,i} \right] - \mathbf{m}_c^{n+1} \\ &= \frac{1}{\gamma \, w_j + \sum_i \alpha_{i,j,c}} \sum_i \alpha_{i,j,c} \left[\mathbf{x}_{j,i} - \mathbf{m}_c^{n+1} \right]\end{aligned}$$

Chapter 6

Conclusions and future work

This chapter presents the main conclusions that can be derived from the work performed and the results obtained in this research. In addition, we propose some interesting lines of future work.

6.1 Conclusions

The lines of work addressed, as well as the interpretation and discussion of the results obtained, allow us to arrive to the following conclusions:

1. We provided a general and updated overview on the basic fundamentals of Magnetic Resonance Imaging methods, the image pre-processing techniques and image analysis methodologies, focusing more deeply in the widely used machine learning algorithms, as well as their advantages and limitations.

2. We implemented the Fisher Map methodology, an supervised learning algorithm, to analyze DWI datasets from the brain of rodents and humans under a fasting paradigm, comparing the results with those obtained from a classical mono-exponential model of diffusion. Both methodologies used the same DWI datasets and yielded consistent results, but the Fisher Maps methodology avoided physiological uncertainties and processing limitations underlying the model-dependent interpretations. Fisher Maps demonstrated the simultaneous activation of hypothalamic and cortico-limbic structures, albeit with different appetite index in rodents and humans. In particular, rodents and humans showed similar hypothalamic activations, but rodents depicted more intense activations of cortico-limbic structures than humans, for similar fasting periods. Finally, the Fisher Maps methodology may be easily applied to other cerebral activation paradigms, complementing conveniently standard analysis protocols.

3. We propose, apply and validate an alternative model-free classification methodology to address the analysis of MRI datasets with small number of subjects, implementing two classification strategies: Pixel Ensemble-Based Classification (PE-BC) and Pixel-Based Classification (P-BC). This methodology was

applied to analyze the DWI datasets of an appetite paradigm and MRI datasets of gliomas, both in rodents and humans. Using the PE-BC strategy, the lowest classification accuracy obtained was 75%, on twelve appetite conditions from six mice, and the highest accuracy was 92% on twelve high-grade gliomas from female and male rat models. The PB-C method was only applied to analyze the appetite paradigm datasets, obtaining 67% of classification accuracy at best. Thus, the PE-BC method achieved the best classification results, can be easily implemented, does not require image registration, and may be extended effortlessly to other image classification problems.

4. We develop and illustrate the performance of a novel unsupervised learning algorithm. This algorithm is able to analyze together data from different sources and group the observations into clusters that may or may not appear in all sources. This procedure allows finding common patterns among all the studied subjects and patterns that are specific to a subset, which is not currently possible by applying standard clustering methods. Our algorithm is a general purpose clustering method, that is to say, although in this work we have focused on its application in experimental data of few subjects, this algorithm can also be applied in other contexts. Another advantage of this algorithm is its flexibility to adapt to the variability of the data in each of the datasets analyzed jointly, being able to optimize different means for the same cluster in different datasets. In this thesis, we have illustrated the advantages of our algorithm by showing the results obtained by analyzing synthetic data and MRI collections from brain tumor studies. The results obtained form the basis for future applications of this algorithm as support in the investigation of brain tumors and other pathologies.

5. In summary, this thesis provides new useful approach for the automatic extraction of relevant information in contexts where small datasets are available. Moreover, novel supervised and unsupervised machine learning methods are proposed, able to work in these contexts.

A detailed discussion about the most revealing results can be found in sections 4.4 and 5.6.

6.2 Future work

Present results pave the way to a plethora of future developments, improving both basic research and diagnostic medicine. The following may be envisioned at present;

1. The Fisher Maps approach may be applied to other cerebral activation paradigms

and non-invasive tests on cerebral performance. Mapping any cerebral stimulation, motor or sensory task, or monitoring the progress of neurodegenerative disorders, constitute reasonably affordable objectives for the near future.

2. Extend PE-BC method with projections that combine adjacent pixels in order to exploit the expected relation of related brain tissues. In this setup bigger datasets will be required to prevent the overfitting caused by the increased complexity of the models.

3. Further improvememts can be added to the multiclustering algorithm. First, the choice of the number of clusters in the presented methods is not automatized. Several methods have been proposed in the scientific literature to accomplish this task[Liu et al., 2010, Sweeney et al., 2015, Tellaroli et al., 2016, Ren et al., 2016] Future work can include this development combining the presented clustering method with techniques such as the validation method based on negentropy [Lago-Fernández et al., 2009, Lago-Fernández and Corbacho, 2010] for determining the optimal number of clusters.

   Another important aspect that can be improved is the ability to allow small variations in the distributions fitted to each of the clusters depending on the data source, specially the covariance matrix. To this end, the likelihood function could be extended allowing transformations on the covariance matrices different in each data source, but including some penalization for too different distribution parameters.

   Finally different distribution mixes could be implemented such as lognormal distribution in order to support other cluster shapes like asymmetric clusters.

4. Apply the proposed clustering algorithm to other pattern recognition problems, for example, to predict the preferences of a user in databases from applications such as Netflix, based on the historical behavior of that user and that of other users related with him.

Chapter 7

Conclusiones y trabajo futuro

Este capítulo presenta las principales conclusiones que pueden derivarse del trabajo realizado y de los resultados obtenidos en esta investigación. Además, proponemos algunas líneas interesantes de trabajo futuro.

7.1 Conclusiones

Las líneas de trabajo abordadas, así como la interpretación y discusión de los resultados obtenidos nos permiten arribar a las siguientes conclusiones:

1. Proporcionamos una descripción general y actualizada de los fundamentos básicos de los métodos de Imagen por Resonancia Magnética (MRI por sus siglas en inglés), las técnicas de pre-procesamiento y los métodos de análisis de imágenes, reseñando en más detalle los algoritmos de *machine learning* más comunmente usados, así como sus principales ventajas y limitaciones.

2. Implementamos la metodología *Fisher Maps*, un algoritmo de aprendizaje supervisado para analizar datasets de imágenes de difusión de cerebros de roedores y humanos bajo estimulación por apetito. Comparamos nuestros resultados con los obtenidos aplicando el clásico modelo mono-exponencial de difusión a los mismos conjuntos de datos; obteniendo resultados consistentes en ambos casos. Sin embargo, la metodología *Fisher Maps* evita la incertidumbre y limitaciones de procesamiento subyacentes en las interpretaciones dependientes de modelo. Los mapas obtenidos demostraron la activación simultánea de estructuras hipotalámicas y cortico-límbicas, aunque con diferentes valores de índice de apetito entre roedores y humanos. En particular, roedores y humanos mostraron similares activaciones hipotalámicas, pero los roedores presentaron activaciones más intensas que los humanos en las estructuras cortico-límbicas para períodos de ayuno similares. Finalmente, la metodología *Fisher Maps* puede aplicarse fácilmente a otros paradigmas de activación cerebral, complementando convenientemente los protocolos de análisis clásicos.

3. Propusimos, aplicamos y validamos una metodología alternativa de clasificación, independiente de modelo, para analizar imágenes MRI con un número limitado de sujetos; implementando dos estrategias de clasificación: *Pixel*

Ensemble-Based Classification (PE-BC) y *Pixel-Based Classification*(P-BC). Esta metodología se empleó para analizar conjuntos de datos de imágenes de MR ponderadas en difusión, adquiridas en roedores y humanos bajo estimulación por apetito; así como también para analizar conjuntos de MRI de tumores en roedores y humanos. Con la estrategia PE-BC la menor precisión de clasificación obtenida fue 75 % para doce condiciones de estímulo por apetito de seis sujetos, y la mayor precisión fue del 92 % para doce gliomas de alto grado en modelos de ratas hembras y machos. Con el método PB-C solo se analizaron los conjuntos de datos de estimulación por apetito, alcanzándose una precisión de clasificación del 67 % en el mejor de los casos. Por tanto, el método PE-BC obtuvo los mejores resultados de clasificación, puede ser facilmente implementado, no requiere alinear las imágenes, y puede ser extendido sin complejidad a otros problemas de clasificación de imágenes.

4. Desarrollamos y mostramos el desempeño de un nuevo algoritmo de aprendizaje no supervisado. Este algoritmo es capaz de analizar conjuntamente datos provenientes de distintas fuentes y agrupar las observaciones en grupos que pueden o no aparecer en todas las fuentes. Este procedimiento permite encontrar patrones comunes a todos los sujetos estudiados y otros específicos a un subconjunto de los mismos, lo cual no es posible actualmente aplicando los métodos de agrupamiento clásicos. Nuestro algoritmo es un método de agrupamiento de propósito general, es decir, aunque en este trabajo nos hemos centrado en su aplicación en datos experimentales de pocos sujetos, este algoritmo puede también ser aplicado en otros contextos. Otra ventaja de este algoritmo es su flexibilidad para adaptarse a la variabilidad de los datos en cada uno de los conjuntos de datos analizados conjuntamente, siendo capaz de optimizar medias diferentes para un mismo grupo en distintos conjuntos de datos. En esta tesis hemos ilustrado las ventajas de nuestro algoritmo mostrando los resultados obtenidos analizando datos sintéticos y colecciones de MRI de estudios de tumores cerebrales. Los resultados obtenidos sientan las bases para futuras aplicaciones de este algoritmo como apoyo en la investigación de tumores cerebrales y otras patologías.

5. En resumen, esta tesis proporciona un nuevo y últil enfoque para la extracción automática de información relevante en contextos donde se dispone de datos provenientes de un número reducido de sujetos. Además, se proponen nuevos métodos de *machine learning*, tanto de aprendizaje supervisado como no supervisado, capaces de trabajar en estos contextos.

Se puede consultar una discusión detallada de los resultados más revelantes en las secciones 4.4 y 5.6.

7.2 Trabajo futuro

Los resultados actuales allanan el camino para otros desarrollos futuros, mejorando tanto la investigación básica como la medicina diagnóstica. En la actualidad se pueden prever los siguientes:

1. La metolología *Fisher Maps* puede aplicarse a otros paradigmas de activación cerebral y pruebas no invasivas sobre el rendimiento cerebral. El mapeo de cualquier estimulación cerebral, la tarea motora o sensorial, o el seguimiento del progreso de los trastornos neurodegenerativos, constituyen objetivos razonablemente asequibles para el futuro próximo.

2. Extender el método *Pixel Ensemble-Based Classification* usando proyecciones que combinen píxeles adyacentes para tener en cuenta la relación entre los tejidos cerebrales relacionados. En esta configuración se requerirán conjuntos de datos más grandes para evitar el sobreajuste causado por la mayor complejidad de los modelos.

3. Implementar mejoras en el algoritmo de agrupamiento presentado. Por un lado, la elección del número de grupos no es automática. En la literatura científica han sido propuestos varios métodos para llevar a cabo esta tarea [Liu et al., 2010, Sweeney et al., 2015, Tellaroli et al., 2016, Ren et al., 2016]. Como trabajo futuro proponemos resolver esta limitación usando técnicas tales como el método de validación basado en negentropía [Lago-Fernández et al., 2009, Lago-Fernández and Corbacho, 2010] para determinar el número óptimo de grupos.

   Otro aspecto importante que se puede mejorar es la capacidad de permitir pequeñas variaciones en las distribuciones ajustadas en cada uno de los grupos dependiendo de la fuente de datos, especialmente en la matriz de covarianza. Para ello, la función de verosimilitud podría extenderse permitiendo transformaciones en las matrices de covarianza diferentes en cada fuente de datos, pero incluyendo algunas penalizaciones para parámetros de distribución demasiado diferentes.

   Finalmente se podrían implementar diferentes mixturas de distribuciones, tales como la distribución lognormal, con el fin de soportar grupos asimétricos y de otras formas.

4. Aplicar el algoritmo de agrupamiento propuesto a otros problemas de reconocimiento de patrones, por ejemplo, para predecir las preferencias de un usuario en bases de datos de aplicaciones como Netflix, basándose en el comportamiento histórico de ese usuario y el de otros usuarios relacionados con él.

Bibliography

[Abragam, 1961] Abragam, A. (1961). *The Principles of Nuclear Magnetism*, volume 1. Oxford University Press.

[Ackerman and Neil, 2010] Ackerman, J. J. and Neil, J. J. (2010). The use of mr-detectable reporter molecules and ions to evaluate diffusion in normal and ischemic brain. *NMR in biomedicine*, 23(7):725–733.

[Ahmed et al., 2014] Ahmed, B., Thesen, T., Blackmon, K., Zhao, Y., Devinsky, O., Kuzniecky, R., and Brodley, C. (2014). Hierarchical conditional random fields for outlier detection: an application to detecting epileptogenic cortical malformations. In *International Conference on Machine Learning*, pages 1080–1088.

[Akaike, 1974] Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723.

[Alexander et al., 2007] Alexander, A. L., Lee, J. E., Lazar, M., and Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics*, 4(3):316–329.

[Allen-Institute, 2017] Allen-Institute (2017). Atlas Brain Map from Allen Institute for Brain Science. http://atlas.brain-map.org/atlas?atlas=138322605#atlas=138322605&plate=112282731&structure=10467&x=40328.53125&y=47069.8671875&zoom=-8&resolution=248.21&z=4. Accessed 2017-05-4.

[Ardekani et al., 2005] Ardekani, B. A., Guckemus, S., Bachman, A., Hoptman, M. J., Wojtaszek, M., and Nierenberg, J. (2005). Quantitative comparison of algorithms for inter-subject registration of 3d volumetric brain mri scans. *Journal of neuroscience methods*, 142(1):67–76.

[Arlot et al., 2010] Arlot, S., Celisse, A., et al. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys*, 4:40–79.

[Armananzas et al., 2016] Armananzas, R., Iglesias, M., Morales, D. A., and Alonso-Nanclares, L. (2016). Voxel-based diagnosis of alzheimer's disease using classifier ensembles. *IEEE journal of biomedical and health informatics*.

[Arnold et al., 2001] Arnold, J. B., Liow, J.-S., Schaper, K. A., Stern, J. J., Sled, J. G., Shattuck, D. W., Worth, A. J., Cohen, M. S., Leahy, R. M., Mazziotta, J. C., et al. (2001). Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects. *NeuroImage*, 13(5):931–943.

[Ashburner, 2009] Ashburner, J. (2009). Computational anatomy with the spm software. *Magnetic resonance imaging*, 27(8):1163–1174.

[Ashburner, 2012] Ashburner, J. (2012). Spm: a history. *Neuroimage*, 62(2):791–800.

[Atalayer et al., 2014] Atalayer, D., Pantazatos, S. P., Gibson, C. D., McOuatt, H., Puma, L., Astbury, N. M., and Geliebter, A. (2014). Sexually dimorphic functional connectivity in response to high vs. low energy-dense food cues in obese humans: An fmri study. *Neuroimage*, 100:405–413.

[Avants et al., 2008] Avants, B. B., Epstein, C. L., Grossman, M., and Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, 12(1):26–41.

[Avants et al., 2011] Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., and Gee, J. C. (2011). A reproducible evaluation of ants similarity metric performance in brain image registration. *Neuroimage*, 54(3):2033–2044.

[Azmi et al., 2013] Azmi, R., Pishgoo, B., Norozi, N., and Yeganeh, S. (2013). Ensemble semi-supervised frame-work for brain magnetic resonance imaging tissue segmentation. *Journal of medical signals and sensors*, 3(2):94.

[Baltzer et al., 2013] Baltzer, P. A., Dietzel, M., and Kaiser, W. A. (2013). A simple and robust classification tree for differentiation between benign and malignant lesions in mr-mammography. *European radiology*, 23(8):2051–2060.

[Bandyopadhyay and Paul, 2013] Bandyopadhyay, S. K. and Paul, T. U. (2013). Segmentation of brain tumour from mri image analysis of k-means and dbscan clustering. *International Journal of Research in Engineering and Science*, 1(1):48–57.

[Banfield and Raftery, 1993] Banfield, J. D. and Raftery, A. E. (1993). Model-based gaussian and non-gaussian clustering. *Biometrics*, pages 803–821.

[Barajas Jr and Cha, 2014] Barajas Jr, R. F. and Cha, S. (2014). Benefits of dynamic susceptibility-weighted contrast-enhanced perfusion mri for glioma diagnosis and therapy. *CNS oncology*, 3(6):407–419.

[Basavanhally et al., 2015] Basavanhally, A., Viswanath, S., and Madabhushi, A. (2015). Predicting classifier performance with limited training data: Applications to computer-aided diagnosis in breast and prostate cancer. *PloS one*, 10(5):e0117900.

[Baselice et al., 2015] Baselice, F., Coppolino, L., D'Antonio, S., Ferraioli, G., and Sgaglione, L. (2015). A dbscan based approach for jointly segment and classify brain mr images. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 2993–2996. IEEE.

[Basser et al., 1994] Basser, P. J., Mattiello, J., and LeBihan, D. (1994). Mr diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259–267.

[Batterham et al., 2007] Batterham, R. L., Rosenthal, J. M., Zelaya, F. O., Barker, G. J., Withers, D. J., Williams, S. C., et al. (2007). Pyy modulation of cortical and hypothalamic brain areas predicts feeding behaviour in humans. *Nature*, 450(7166):106–109.

[Behroozi and Daliri, 2014] Behroozi, M. and Daliri, M. R. (2014). Predicting brain states associated with object categories from fmri data. *Journal of integrative neuroscience*, 13(04):645–667.

[Belaroussi et al., 2006] Belaroussi, B., Milles, J., Carme, S., Zhu, Y. M., and Benoit-Cattin, H. (2006). Intensity non-uniformity correction in mri: existing methods and their validation. *Medical Image Analysis*, 10(2):234–246.

[Belkin and Niyogi, 2003] Belkin, M. and Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396.

[Benda et al., 1971] Benda, P., Someda, K., Messer, J., and Sweet, W. H. (1971). Morphological and immunochemical studies of rat glial tumors and clonal strains propagated in culture. *Journal of neurosurgery*, 34(3):310–323.

[Bengio et al., 2007] Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H., et al. (2007). Greedy layer-wise training of deep networks. *Advances in neural information processing systems*, 19:153.

[Berthoud and Morrison, 2008] Berthoud, H.-R. and Morrison, C. (2008). The brain, appetite, and obesity. *Annu. Rev. Psychol.*, 59:55–92.

[Besson et al., 2008] Besson, P., Bernasconi, N., Colliot, O., Evans, A., and Bernasconi, A. (2008). Surface-based texture and morphological analysis detects subtle cortical dysplasia. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, pages 645–652.

[Bezdek et al., 1984] Bezdek, J. C., Ehrlich, R., and Full, W. (1984). Fcm: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, 10(2-3):191–203.

[Bilmes et al., 1998] Bilmes, J. A. et al. (1998). A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. *International Computer Science Institute*, 4(510):126.

[Bishop, 2006] Bishop, C. M. (2006). Pattern recognition. *Machine Learning*, 128:1–58.

[Blair and Nichaman, 2002] Blair, S. N. and Nichaman, M. Z. (2002). The public health problem of increasing prevalence rates of obesity and what should be done about it. In *Mayo Clinic Proceedings*, volume 77, pages 109–113. Elsevier.

[Bloch, 1946] Bloch, F. (1946). Nuclear induction. *Physical review*, 70(7-8):460.

[Bojorquez et al., 2016] Bojorquez, J. Z., Bricq, S., Brunotte, F., Walker, P. M., and Lalande, A. (2016). A novel alternative to classify tissues from t1 and t2 relaxation times for prostate mri. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 29(5):777–788.

[Boyes et al., 2008] Boyes, R. G., Gunter, J. L., Frost, C., Janke, A. L., Yeatman, T., Hill, D. L., Bernstein, M. A., Thompson, P. M., Weiner, M. W., Schuff, N., et al. (2008). Intensity non-uniformity correction using n3 on 3-t scanners with multichannel phased array coils. *Neuroimage*, 39(4):1752–1762.

[Brada, 2006] Brada, M. (2006). Radiotherapy in malignant glioma. *Annals of Oncology*, 17(suppl_10):x183.

[Bradley et al., 1999] Bradley, W., Chen, D., and Atkinson, D. (1999). Fast spin-echo and echo-planar imaging. In Stark, D. and Bradley, W., editors, *Magnetic resonance imaging*, pages 125–158. Mosby: St. Louis, Missouri, 3rd edition.

[Brady et al., 1985] Brady, T., Wismer, G., Buxton, R., Stark, D., and Rosen, B. (1985). Magnetic resonance chemical shift imaging. *Magnetic resonance annual*, pages 55–80.

[Brateman, 1986] Brateman, L. (1986). Chemical shift imaging: a review. *American Journal of Roentgenology*, 146(5):971–980.

[Breiman, 1996] Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2):123–140.

[Breiman, 2001] Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

[Brink et al., 1989] Brink, H. F., Buschmann, M. D., and Rosen, B. R. (1989). Nmr chemical shift imaging. *Computerized Medical Imaging and Graphics*, 13(1):93–104.

[Brosch et al., 2016] Brosch, T., Tang, L. Y., Yoo, Y., Li, D. K., Traboulsee, A., and Tam, R. (2016). Deep 3d convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation. *IEEE transactions on medical imaging*, 35(5):1229–1239.

[Bullmore et al., 1999] Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., and Brammer, M. J. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural mr images of the brain. *IEEE transactions on medical imaging*, 18(1):32–42.

[Burges, 1998] Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167.

[Bydder et al., 1998] Bydder, G., Hajnal, J., and Young, I. (1998). Mri: use of the inversion recovery pulse sequence. *Clinical radiology*, 53(3):159–176.

[Cacha et al., 2016] Cacha, L., Parida, S., Dehuri, S., Cho, S.-B., and Poznanski, R. (2016). A fuzzy integral method based on the ensemble of neural networks to analyze fmri data for cognitive state classification across multiple subjects. *Journal of integrative neuroscience*, 15(04):593–606.

[Caiazzo et al., 2016] Caiazzo, G., Trojsi, F., Cirillo, M., Tedeschi, G., and Esposito, F. (2016). Q-ball imaging models: comparison between high and low angular resolution diffusion-weighted mri protocols for investigation of brain white matter integrity. *Neuroradiology*, 58(2):209–215.

[Caliński and Harabasz, 1974] Caliński, T. and Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, 3(1):1–27.

[Cao et al., 2016] Cao, W., Liang, C., Gen, Y., Wang, C., Zhao, C., and Sun, L. (2016). Role of diffusion-weighted imaging for detecting bone marrow infiltration in skull in children with acute lymphoblastic leukemia. *Diagnostic and Interventional Radiology*, 22(6):580.

[Carnell et al., 2012] Carnell, S., Gibson, C., Benson, L., Ochner, C., and Geliebter, A. (2012). Neuroimaging and obesity: current knowledge and future directions. *Obesity Reviews*, 13(1):43–56.

[Cea et al., 2012] Cea, V., Sala, C., and Verpelli, C. (2012). Antiangiogenic therapy for glioma. *Journal of signal transduction*, 2012.

[Chaddad and Tanougast, 2016] Chaddad, A. and Tanougast, C. (2016). Quantitative evaluation of robust skull stripping and tumor detection applied to axial mr images. *Brain Informatics*, 3(1):53–61.

[Chang et al., 2017] Chang, H., Huang, W., Wu, C., Huang, S., Guan, C., Sekar, S., Bhakoo, K. K., and Duan, Y. (2017). A new variational method for bias correction and its applications to rodent brain extraction. *IEEE transactions on medical imaging*, 36(3):721–733.

[Chang et al., 2012] Chang, Y.-C., Huang, Y.-H., Huang, C.-S., Chang, P.-K., Chen, J.-H., and Chang, R.-F. (2012). Classification of breast mass lesions using model-based analysis of the characteristic kinetic curve derived from fuzzy c-means clustering. *Magnetic resonance imaging*, 30(3):312–322.

[Chavhan et al., 2009] Chavhan, G. B., Babyn, P. S., Thomas, B., Shroff, M. M., and Haacke, E. M. (2009). Principles, techniques, and applications of t2*-based mr imaging and its special applications 1. *Radiographics*, 29(5):1433–1449.

[Chou et al., 2011] Chou, N., Wu, J., Bingren, J. B., Qiu, A., and Chuang, K.-H. (2011). Robust automatic rodent brain extraction using 3-d pulse-coupled neural networks (pcnn). *IEEE Transactions on Image Processing*, 20(9):2554–2564.

[Churchill et al., 2012a] Churchill, N. W., Oder, A., Abdi, H., Tam, F., Lee, W., Thomas, C., Ween, J. E., Graham, S. J., and Strother, S. C. (2012a). Optimizing preprocessing and analysis pipelines for single-subject fmri. i. standard temporal motion and physiological noise correction methods. *Human brain mapping*, 33(3):609–627.

[Churchill et al., 2017] Churchill, N. W., Raamana, P. R., Spring, R., and Strother, S. C. (2017). Optimizing fmri preprocessing pipelines for block-design tasks as a function of age. *NeuroImage*.

[Churchill et al., 2015] Churchill, N. W., Spring, R., Afshin-Pour, B., Dong, F., and Strother, S. C. (2015). An automated, adaptive framework for optimizing preprocessing pipelines in task-based functional mri. *PloS one*, 10(7):e0131520.

[Churchill et al., 2012b] Churchill, N. W., Yourganov, G., Oder, A., Tam, F., Graham, S. J., and Strother, S. C. (2012b). Optimizing preprocessing and analysis pipelines for single-subject fmri: 2. interactions with ica, pca, task contrast and inter-subject heterogeneity. *PloS one*, 7(2):e31147.

[Clark et al., 2013] Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., et al. (2013). The cancer imaging archive (tcia): maintaining and operating a public information repository. *Journal of digital imaging*, 26(6):1045–1057.

[Coifman et al., 2005a] Coifman, R. R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F., and Zucker, S. W. (2005a). Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the National Academy of Sciences of the United States of America*, 102(21):7426–7431.

[Coifman et al., 2005b] Coifman, R. R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F., and Zucker, S. W. (2005b). Geometric diffusions as a tool for harmonic analysis and structure definition of data: Multiscale methods. *Proceedings of the National Academy of Sciences of the United States of America*, 102(21):7432–7437.

[Coll et al., 2007] Coll, A. P., Farooqi, I. S., and O'Rahilly, S. (2007). The hormonal control of food intake. *Cell*, 129(2):251–262.

[Coll and Yeo, 2013] Coll, A. P. and Yeo, G. S. (2013). The hypothalamus and metabolism: integrating signals to control energy and glucose homeostasis. *Current opinion in pharmacology*, 13(6):970–976.

[Comon, 1994] Comon, P. (1994). Independent component analysis, a new concept? *Signal processing*, 36(3):287–314.

[Conn et al., 2000] Conn, A. R., Gould, N. I., and Toint, P. L. (2000). Nonlinear equations and nonlinear fitting. In *Trust region methods*, pages 749–774. SIAM.

[Cover and Hart, 1967] Cover, T. and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27.

[Cox, 1996] Cox, R. W. (1996). Afni: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*, 29(3):162–173.

[Cox, 2012] Cox, R. W. (2012). Afni: what a long strange trip it's been. *Neuroimage*, 62(2):743–747.

[Cuingnet et al., 2011] Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M.-O., Chupin, M., Benali, H., Colliot, O., Initiative, A. D. N., et al. (2011). Automatic classification of patients with alzheimer's disease from structural mri: a comparison of ten methods using the adni database. *neuroimage*, 56(2):766–781.

[Dale et al., 1999] Dale, A. M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis: I. segmentation and surface reconstruction. *Neuroimage*, 9(2):179–194.

[Damadian et al., 1974] Damadian, R., Zaner, K., Hor, D., and DiMaio, T. (1974). Human tumors detected by nuclear magnetic resonance. *Proceedings of the National Academy of Sciences*, 71(4):1471–1473.

[Davies and Bouldin, 1979] Davies, D. L. and Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1(2):224–227.

[de Bruijne, 2016] de Bruijne, M. (2016). Machine learning approaches in medical image analysis: From detection to diagnosis. *Medical Image Analysis*, 33:94 – 97. 20th anniversary of the Medical Image Analysis journal (MedIA).

[De La Rocha et al., 2014] De La Rocha, A., Sampron, N., Alonso, M. M., and Matheu, A. (2014). Role of sox family of transcription factors in central nervous system tumors. *Am J Cancer Res*, 4(4):312–324.

[De Nunzio et al., 2015] De Nunzio, G., Cataldo, R., and Carlà, A. (2015). Robust intensity standardization in brain magnetic resonance images. *Journal of digital imaging*, 28(6):727–737.

[De Silva et al., 2012] De Silva, A., Salem, V., Matthews, P. M., and Dhillo, W. S. (2012). The use of functional mri to study appetite control in the cns. *Experimental diabetes research*, 2012.

[DeBerardinis et al., 2008] DeBerardinis, R. J., Lum, J. J., Hatzivassiliou, G., and Thompson, C. B. (2008). The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell metabolism*, 7(1):11–20.

[Demir and Cetingül, 2015] Demir, A. and Cetingül, H. E. (2015). Sequential hierarchical agglomerative clustering of white matter fiber pathways. *IEEE Transactions on Biomedical Engineering*, 62(6):1478–1489.

[Demir et al., 2013] Demir, A., Mohamed, A., and Çetingül, H. E. (2013). Online agglomerative hierarchical clustering of neural fiber tracts. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 85–88. IEEE.

[Dempster et al., 1977] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38.

[Dhermain, 2014] Dhermain, F. (2014). Radiotherapy of high-grade gliomas: current standards and new concepts, innovations in imaging and radiotherapy, and new therapeutic approaches. *Chin J Cancer*, 33(1):16–24.

[Dietzel et al., 2012] Dietzel, M., Baltzer, P. A., Dietzel, A., Zoubi, R., Gröschel, T., Burmeister, H. P., Bogdan, M., and Kaiser, W. A. (2012). Artificial neural networks for differential diagnosis of breast lesions in mr-mammography: A systematic approach addressing the influence of network architecture on diagnostic performance using a large clinical database. *European journal of radiology*, 81(7):1508–1513.

[Dikaios et al., 2015a] Dikaios, N., Alkalbani, J., Abd-Alazeez, M., Sidhu, H. S., Kirkham, A., Ahmed, H. U., Emberton, M., Freeman, A., Halligan, S., Taylor, S., et al. (2015a). Zone-specific logistic regression models improve classification of prostate cancer on multi-parametric mri. *European radiology*, 25(9):2727–2737.

[Dikaios et al., 2015b] Dikaios, N., Alkalbani, J., Sidhu, H. S., Fujiwara, T., Abd-Alazeez, M., Kirkham, A., Allen, C., Ahmed, H., Emberton, M., Freeman, A., et al. (2015b). Logistic regression model for diagnosis of transition zone prostate cancer on multi-parametric mri. *European radiology*, 25(2):523–532.

[Domingos and Pazzani, 1997] Domingos, P. and Pazzani, M. (1997). On the optimality of the simple bayesian classifier under zero-one loss. *Machine learning*, 29(2):103–130.

[Donoso et al., 2010] Donoso, R., Veloz, A., and Allende, H. (2010). Modified expectation maximization algorithm for mri segmentation. In *Iberoamerican Congress on Pattern Recognition*, pages 63–70. Springer.

[Dou et al., 2016] Dou, Q., Chen, H., Yu, L., Zhao, L., Qin, J., Wang, D., Mok, V. C., Shi, L., and Heng, P.-A. (2016). Automatic detection of cerebral microbleeds from mr images via 3d convolutional neural networks. *IEEE transactions on medical imaging*, 35(5):1182–1195.

[Duda et al., 2012] Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

[Dunn, 1974] Dunn, J. C. (1974). Well-separated clusters and optimal fuzzy partitions. *Journal of cybernetics*, 4(1):95–104.

[Eklund et al., 2012] Eklund, A., Andersson, M., Josephson, C., Johannesson, M., and Knutsson, H. (2012). Does parametric fmri analysis with spm yield valid results?—an empirical study of 1484 rest datasets. *NeuroImage*, 61(3):565–578.

[Eklund et al., 2016] Eklund, A., Nichols, T. E., and Knutsson, H. (2016). Cluster failure: why fmri inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, page 201602413.

[El Azami et al., 2013] El Azami, M., Hammers, A., Costes, N., and Lartizien, C. (2013). Computer aided diagnosis of intractable epilepsy with mri imaging based on textural information. In *Pattern Recognition in Neuroimaging (PRNI), 2013 International Workshop on*, pages 90–93. IEEE.

[El-Dahshan et al., 2010] El-Dahshan, E.-S. A., Hosny, T., and Salem, A.-B. M. (2010). Hybrid intelligent techniques for mri brain images classification. *Digital Signal Processing*, 20(2):433–441.

[El-Dahshan et al., 2014] El-Dahshan, E.-S. A., Mohsen, H. M., Revett, K., and Salem, A.-B. M. (2014). Computer-aided diagnosis of human brain tumor through mri: A survey and a new algorithm. *Expert systems with Applications*, 41(11):5526–5545.

[Elazab et al., 2015] Elazab, A., Wang, C., Jia, F., Wu, J., Li, G., and Hu, Q. (2015). Segmentation of brain tissues from magnetic resonance images using adaptively regularized kernel-based fuzzy-means clustering. *Computational and mathematical methods in medicine*, 2015.

[Ely et al., 2014] Ely, A. V., Childress, A. R., Jagannathan, K., and Lowe, M. R. (2014). Differential reward response to palatable food cues in past and current dieters: a fmri study. *Obesity*, 22(5).

[Erickson et al., 2017] Erickson, B. J., Korfiatis, P., Akkus, Z., and Kline, T. L. (2017). Machine learning for medical imaging. *RadioGraphics*, 37(2):505–515.

[Ernst et al., 1987] Ernst, R. R., Bodenhausen, G., Wokaun, A., et al. (1987). *Principles of nuclear magnetic resonance in one and two dimensions*, volume 14. Clarendon Press Oxford.

[Ester et al., 1996] Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231.

[Fellah et al., 2012] Fellah, S., Callot, V., Viout, P., Confort-Gouny, S., Scavarda, D., Dory-Lautrec, P., Figarella-Branger, D., Cozzone, P. J., and Girard, N. (2012). Epileptogenic brain lesions in children: the added-value of combined diffusion imaging and proton mr spectroscopy to the presurgical differential diagnosis. *Child's Nervous System*, 28(2):273–282.

[Fennema-Notestine et al., 2006] Fennema-Notestine, C., Ozyurt, I. B., Clark, C. P., Morris, S., Bischoff-Grethe, A., Bondi, M. W., Jernigan, T. L., Fischl, B., Segonne, F., Shattuck, D. W., et al. (2006). Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: Effects of diagnosis, bias correction, and slice location. *Human brain mapping*, 27(2):99–113.

[Fisher, 1935] Fisher, R. A. (1935). The design of experiments. 1935. *Oliver and Boyd, Edinburgh*.

[Fisher, 1936] Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188.

[FNLCR, 2014] FNLCR (2014). The Cancer Imaging Archive (TCIA) Public Access. https://wiki.cancerimagingarchive.net/display/Public/Wiki#collections-list. Accessed 2017-05-10.

[Fontana et al., 2016] Fontana, G., Riboldi, M., Gianoli, C., Chirvase, C. I., Villa, G., Paganelli, C., Summers, P. E., Tagaste, B., Pella, A., Fossati, P., et al. (2016). Mri quantification of pancreas motion as a function of patient setup for particle therapy—a preliminary study. *Journal of Applied Clinical Medical Physics*, 17(5).

[Franke et al., 2010] Franke, K., Ziegler, G., Klöppel, S., Gaser, C., Initiative, A. D. N., et al. (2010). Estimating the age of healthy subjects from t 1-weighted mri scans using kernel methods: Exploring the influence of various parameters. *Neuroimage*, 50(3):883–892.

[Freund and Schapire, 1995] Freund, Y. and Schapire, R. E. (1995). A desicion-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory*, pages 23–37. Springer.

[Friedman et al., 2009] Friedman, J., Hastie, T., and Tibshirani, R. (2009). *The elements of statistical learning*. Springer series in statistics Springer, second edition.

[Friston et al., 2007] Friston, K. J., Ashburner, J., Kiebel, S. J., Nichols, T., and Penny, W. (2007). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press.

[Friston et al., 1994] Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping*, 2(4):189–210.

[Fulton, 2009] Fulton, S. (2009). Neurosystems linking corticolimbic and hypothalamic pathways in energy balance: view from the chair. *International Journal of Obesity*, 33:S3–S7.

[Galla et al., 2017] Galla, N., Chiang, G., Chakraborty, S., Singh, R., Tsiouris, A. J., Boockvar, J., and Kovanlikaya, I. (2017). Apparent diffusion coefficient changes predict survival after intra-arterial bevacizumab treatment in recurrent glioblastoma. *Neuroradiology*, 59(5):499–505.

[Gandjbakhch et al., 2013] Gandjbakhch, F., Haavardsholm, E. A., Conaghan, P. G., Ejbjerg, B., Foltz, V., Brown, A. K., Døhn, U. M., Lassere, M., Freeston, J. E., Olsen, I. C., et al. (2013). Determining a magnetic resonance imaging inflammatory activity acceptable state without subsequent radiographic progression in rheumatoid arthritis: results from a followup mri study of 254 patients in clinical remission or low disease activity. *The Journal of rheumatology*, pages jrheum–131088.

[Garey et al., 1982] Garey, M., Johnson, D., and Witsenhausen, H. (1982). The complexity of the generalized lloyd-max problem (corresp.). *IEEE Transactions on Information Theory*, 28(2):255–256.

[Gautier et al., 2000] Gautier, J.-F., Chen, K., Salbe, A. D., Bandy, D., Pratley, R. E., Heiman, M., Ravussin, E., Reiman, E. M., and Tataranni, P. A. (2000). Differential brain responses to satiation in obese and lean men. *Diabetes*, 49(5):838–846.

[Ge et al., 2000] Ge, Y., Udupa, J. K., Nyul, L. G., Wei, L., and Grossman, R. I. (2000). Numerical tissue characterization in ms via standardization of the mr image intensity scale. *Journal of Magnetic Resonance Imaging*, 12(5):715–721.

[Geliebter et al., 2013] Geliebter, A., Pantazatos, S. P., McOuatt, H., Puma, L., Gibson, C. D., and Atalayer, D. (2013). Sex-based fmri differences in obese humans in response to high vs. low energy food cues. *Behavioural brain research*, 243:91–96.

[Gibbons and Chakraborti, 2011] Gibbons, J. D. and Chakraborti, S. (2011). The wilcoxon rank-sum test and confidence interval. In *Nonparametric statistical inference*, pages 290–293. Springer.

[Gibson et al., 2010] Gibson, E., Gao, F., Black, S. E., and Lobaugh, N. J. (2010). Automatic segmentation of white matter hyperintensities in the elderly using flair images at 3t. *Journal of Magnetic Resonance Imaging*, 31(6):1311–1322.

[Glasser et al., 2013] Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., et al. (2013). The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, 80:105–124.

[Glorot and Bengio, 2010] Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Aistats*, volume 9, pages 249–256.

[Gore and Kennan, 1999] Gore, J. C. and Kennan, R. P. (1999). Physical and physiological basis of magnetic relaxation. *Magn. Reson. Imaging*, 1:33–42.

[Graña et al., 2011] Graña, M., Termenon, M., Savio, A., Gonzalez-Pinto, A., Echeveste, J., Pérez, J., and Besga, A. (2011). Computer aided diagnosis system for alzheimer disease using brain diffusion tensor imaging features selected by pearson's correlation. *Neuroscience letters*, 502(3):225–229.

[Gray et al., 2011] Gray, C., MacGillivray, T. J., Eeley, C., Stephens, N. A., Beggs, I., Fearon, K. C., and Greig, C. A. (2011). Magnetic resonance imaging with k-means clustering objectively measures whole muscle volume compartments in sarcopenia/cancer cachexia. *Clinical Nutrition*, 30(1):106–111.

[Griffis et al., 2016] Griffis, J. C., Allendorfer, J. B., and Szaflarski, J. P. (2016). Voxel-based gaussian naïve bayes classification of ischemic stroke lesions in individual t1-weighted mri scans. *Journal of neuroscience methods*, 257:97–108.

[Guo et al., 2014] Guo, Y., Gao, Y., Shao, Y., Price, T., Oto, A., and Shen, D. (2014). Deformable segmentation of 3d mr prostate images via distributed discriminative dictionary and ensemble learning. *Medical physics*, 41(7).

[Guyader et al., 2015] Guyader, J.-M., Bernardin, L., Douglas, N. H., Poot, D. H., Niessen, W. J., and Klein, S. (2015). Influence of image registration on apparent diffusion coefficient images computed from free-breathing diffusion mr images of the abdomen. *Journal of Magnetic Resonance Imaging*, 42(2):315–330.

[Guyon and Elisseeff, 2003] Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182.

[Habes et al., 2013] Habes, I., Krall, S. C., Johnston, S., Yuen, K., Healy, D., Goebel, R., Sorger, B., and Linden, D. E. J. (2013). Pattern classification of valence in depression. *NeuroImage: clinical*, 2:675–683.

[Hahn and Peitgen, 2000] Hahn, H. K. and Peitgen, H.-O. (2000). The skull stripping problem in mri solved by a single 3d watershed transform. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 134–143. Springer.

[Halkidi et al., 2001] Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2001). On clustering validation techniques. *Journal of intelligent information systems*, 17(2):107–145.

[Halkidi et al., 2002a] Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2002a). Cluster validity methods: part i. *ACM Sigmod Record*, 31(2):40–45.

[Halkidi et al., 2002b] Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2002b). Clustering validity checking methods: part ii. *ACM Sigmod Record*, 31(3):19–27.

[Hand and Yu, 2001] Hand, D. J. and Yu, K. (2001). Idiot's bayes—not so stupid after all? *International statistical review*, 69(3):385–398.

[Hartley et al., 2006] Hartley, S., Scher, A., Korf, E., White, L., and Launer, L. (2006). Analysis and validation of automated skull stripping tools: a validation study based on 296 mr images from the honolulu asia aging study. *NeuroImage*, 30(4):1179–1186.

[Hashemi et al., 2012] Hashemi, R. H., Bradley, W. G., and Lisanti, C. J. (2012). *MRI: the basics*. Lippincott Williams & Wilkins.

[Hata et al., 2005] Hata, Y., Kobashi, S., Kondo, K., Kitamura, Y. T., and Yanagida, T. (2005). Transcranial ultrasonography system for visualizing skull and brain surface aided by fuzzy expert system. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 35(6):1360–1373.

[Hendrick and Raff, 1992] Hendrick, R. E. and Raff, U. (1992). Image contrast and noise. *Magnetic resonance imaging*, 1:109–144.

[Hennig et al., 1986] Hennig, J., Nauerth, A., and Friedburg, H. (1986). Rare imaging: a fast imaging method for clinical mr. *Magnetic resonance in medicine*, 3(6):823–833.

[Hill et al., 2001] Hill, D. L., Batchelor, P. G., Holden, M., and Hawkes, D. J. (2001). Medical image registration. *Physics in medicine and biology*, 46(3):R1.

[Hirose et al., 2015] Hirose, S., Nambu, I., and Naito, E. (2015). An empirical solution for over-pruning with a novel ensemble-learning method for fmri decoding. *Journal of neuroscience methods*, 239:238–245.

[Hori et al., 2012] Hori, M., Fukunaga, I., Masutani, Y., Taoka, T., Kamagata, K., Suzuki, Y., and Aoki, S. (2012). Visualizing non-gaussian diffusion: Clinical application of q-space imaging and diffusional kurtosis imaging of the brain and spine. *Magnetic Resonance in Medical Sciences*, 11(4):221–233.

[Hou, 2006] Hou, Z. (2006). A review on mr image intensity inhomogeneity correction. *International Journal of Biomedical Imaging*, 2006.

[Huang et al., 2015] Huang, Y., Chen, X., Zhang, Z., Yan, L., Pan, D., Liang, C., and Liu, Z. (2015). Mri quantification of non-gaussian water diffusion in normal human kidney: a diffusional kurtosis imaging study. *NMR in Biomedicine*, 28(2):154–161.

[Hutterer et al., 2014] Hutterer, M., Hattingen, E., Palm, C., Proescholdt, M. A., and Hau, P. (2014). Current standards and new concepts in mri and pet response assessment of antiangiogenic therapies in high-grade glioma patients. *Neuro-oncology*, page nou322.

[Hutton and Braun, 2003] Hutton, B. F. and Braun, M. (2003). Software for image registration: algorithms, accuracy, efficacy. In *Seminars in nuclear medicine*, volume 33, pages 180–192. Elsevier.

[Jenkinson et al., 2012] Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., and Smith, S. M. (2012). Fsl. *Neuroimage*, 62(2):782–790.

[Jenkinson et al., 2005] Jenkinson, M., Pechaud, M., and Smith, S. (2005). Bet2: Mr-based estimation of brain, skull and scalp surfaces. In *Eleventh annual meeting of the organization for human brain mapping*, volume 17, page 167. Toronto.

[Jensen et al., 2016] Jensen, J. H., Glenn, G. R., and Helpern, J. A. (2016). Fiber ball imaging. *NeuroImage*, 124:824–833.

[Jensen et al., 2017] Jensen, J. H., McKinnon, E. T., Glenn, G. R., and Helpern, J. A. (2017). Evaluating kurtosis-based diffusion mri tissue models for white matter with fiber ball imaging. *NMR in Biomedicine*, 30(5).

[Johnson and Fischer, 2015] Johnson, J. E. and Fischer, K. J. (2015). Results of automatic image registration are dependent on initial manual registration. *Computer methods in biomechanics and biomedical engineering*, 18(16):1856–1861.

[Jolliffe, 2002] Jolliffe, I. (2002). *Principal component analysis*. Wiley Online Library.

[Jung et al., 2014] Jung, S. C., Yeom, J., Kim, J.-H., Ryoo, I., Kim, S., Shin, H., Lee, A., Yun, T., Park, C.-K., Sohn, C.-H., et al. (2014). Glioma: application of histogram analysis of pharmacokinetic parameters from t1-weighted dynamic contrast-enhanced mr imaging to tumor grading. *American Journal of Neuroradiology*, 35(6):1103–1110.

[Kahn et al., 2006] Kahn, S. E., Hull, R. L., and Utzschneider, K. M. (2006). Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature*, 444(7121):840–846.

[Kalavathi and Prasath, 2016] Kalavathi, P. and Prasath, V. S. (2016). Methods on skull stripping of mri head scan images—a review. *Journal of digital imaging*, 29(3):365–379.

[Kaya et al., 2017] Kaya, I. E., Pehlivanlı, A. Ç., Sekizkardeş, E. G., and Ibrikci, T. (2017). Pca based clustering for brain tumor segmentation of t1w mri images. *Computer Methods and Programs in Biomedicine*, 140:19–28.

[Kazerooni et al., 2016] Kazerooni, A. F., Malek, M., Haghighatkhah, H., Parviz, S., Nabil, M., Torbati, L., Assili, S., Saligheh Rad, H., and Gity, M. (2016). Semi-quantitative dynamic contrast-enhanced mri for accurate classification of complex adnexal masses. *Journal of Magnetic Resonance Imaging*.

[Ke et al., 2014] Ke, C., Tran, K., Chen, Y., Di Donato, A. T., Yu, L., Hu, Y., Linskey, M. E., Wang, P. H., Limoli, C. L., and Zhou, Y.-H. (2014). Linking differential radiation responses to glioma heterogeneity. *Oncotarget*, 5(6):1657–1665.

[Keszei et al., 2017] Keszei, A. P., Berkels, B., and Deserno, T. M. (2017). Survey of non-rigid registration tools in medicine. *Journal of digital imaging*, 30(1):102–116.

[Khasraw et al., 2014] Khasraw, M., Ameratunga, M. S., Grant, R., Wheeler, H., and Pavlakis, N. (2014). Antiangiogenic therapy for high-grade glioma. *The Cochrane Library*.

[Khedher et al., 2015] Khedher, L., Ramírez, J., Górriz, J. M., and Brahim, A. (2015). Automatic classification of segmented mri data combining independent component analysis and support vector machines. *Innovation in Medicine and Healthcare 2014*, 207:271.

[Kilkenny et al., 2010] Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M., and Altman, D. G. (2010). Improving bioscience research reporting: the arrive guidelines for reporting animal research. *PLoS Biol*, 8(6):e1000412.

[Kim and Park, 2017] Kim, K. H. and Park, S.-H. (2017). Artificial neural network for suppression of banding artifacts in balanced steady-state free precession mri. *Magnetic Resonance Imaging*, 37:139–146.

[Klein et al., 2009] Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G. E., Collins, D. L., Gee, J., Hellier, P., et al. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *Neuroimage*, 46(3):786–802.

[Kocevar et al., 2016] Kocevar, G., Stamile, C., Hannoun, S., Cotton, F., Vukusic, S., Durand-Dubief, F., and Sappey-Marinier, D. (2016). Graph theory-based brain connectivity for automatic classification of multiple sclerosis clinical courses. *Frontiers in neuroscience*, 10.

[Kopelman, 2000] Kopelman, P. G. (2000). Obesity as a medical problem. *Nature*, 404(6778):635–643.

[Kriegel et al., 2011] Kriegel, H.-P., Kröger, P., Sander, J., and Zimek, A. (2011). Density-based clustering. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(3):231–240.

[Kuhn and Johnson, 2013] Kuhn, M. and Johnson, K. (2013). An introduction to feature selection. In *Applied Predictive Modeling*, pages 487–519. Springer.

[Kuo et al., 2006] Kuo, Y.-T., Herlihy, A. H., So, P.-W., and Bell, J. D. (2006). Manganese-enhanced magnetic resonance imaging (memri) without compromise of the blood–brain barrier detects hypothalamic neuronal activity in vivo. *NMR in biomedicine*, 19(8):1028–1034.

[Lago-Fernández et al., 2009] Lago-Fernández, L., Sánchez-Montañés, M., and Corbacho, F. (2009). Fuzzy cluster validation using the partition negentropy criterion. *Artificial Neural Networks–ICANN 2009*, pages 235–244.

[Lago-Fernández and Corbacho, 2010] Lago-Fernández, L. F. and Corbacho, F. (2010). Normality-based validation for crisp clustering. *Pattern Recognition*, 43(3):782–795.

[Larsen and Aone, 1999] Larsen, B. and Aone, C. (1999). Fast and effective text mining using linear-time document clustering. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 16–22. ACM.

[Lauterbur, 1973] Lauterbur, P. C. (1973). Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature*, 242(5394):190–191.

[Le Bihan, 2003] Le Bihan, D. (2003). Looking into the functional architecture of the brain with diffusion mri. *Nature Reviews Neuroscience*, 4(6):469–480.

[Le Bihan, 2007] Le Bihan, D. (2007). The 'wet mind': water and functional neuroimaging. *Physics in medicine and biology*, 52(7):R57.

[Le Bihan, 2012] Le Bihan, D. (2012). Diffusion, confusion and functional mri. *Neuroimage*, 62(2):1131–1136.

[Le Bihan et al., 1986] Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., and Laval-Jeantet, M. (1986). Mr imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2):401–407.

[Le Bihan et al., 1993] Le Bihan, D., Douek, P., Argyropoulou, M., Turner, R., Patronas, N., and Fulham, M. (1993). Diffusion and perfusion magnetic resonance imaging in brain tumors. *Topics in Magnetic Resonance Imaging*, 5(2):25–31.

[Le Bihan and Iima, 2015] Le Bihan, D. and Iima, M. (2015). Diffusion magnetic resonance imaging: what water tells us about biological tissues. *PLoS Biol*, 13(7):e1002203.

[Le Cun et al., 1990] Le Cun, Y., Jackel, L., Boser, B., Denker, J., Graf, H., Guyon, I., Henderson, D., Howard, R., and Hubbard, W. (1990). Handwritten digit recognition: Applications of neural net chips and automatic learning. In *Neurocomputing*, pages 303–318. Springer.

[LeCun, 2012] LeCun, Y. (2012). Learning invariant feature hierarchies. In *Computer vision–ECCV 2012. Workshops and demonstrations*, pages 496–505. Springer.

[LeCun et al., 2015] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.

[LeCun et al., 1998] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

[Lee et al., 2016] Lee, M.-C., Chuang, K.-S., Chen, M.-K., Liu, C.-K., Lee, K.-W., Tsai, H.-Y., and Lin, H.-H. (2016). Fuzzy c-means clustering of magnetic resonance imaging on apparent diffusion coefficient maps for predicting nodal metastasis in head and neck cancer. *The British journal of radiology*, 89(1063):20150059.

[Li et al., 2013] Li, J., Liu, X., Zhuo, J., Gullapalli, R. P., and Zara, J. M. (2013). An automatic rat brain extraction method based on a deformable surface model. *Journal of neuroscience methods*, 218(1):72–82.

[Liberman et al., 2013] Liberman, G., Louzoun, Y., Aizenstein, O., Blumenthal, D. T., Bokstein, F., Palmon, M., Corn, B. W., and Bashat, D. B. (2013). Automatic multi-modal mr tissue classification for the assessment of response to bevacizumab in patients with glioblastoma. *European journal of radiology*, 82(2):e87–e94.

[Liu et al., 2010] Liu, Y., Li, Z., Xiong, H., Gao, X., and Wu, J. (2010). Understanding of internal clustering validation measures. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on*, pages 911–916. IEEE.

[Lizarbe et al., 2013] Lizarbe, B., Benítez, A., Sánchez-Montañés, M., Lago-Fernández, L. F., Garcia-Martin, M. L., López-Larrubia, P., and Cerdán, S. (2013). Imaging hypothalamic activity using diffusion weighted magnetic resonance imaging in the mouse and human brain. *Neuroimage*, 64:448–457.

[Lizarbe et al., 2015] Lizarbe, B., López-Larrubia, P., and Cerdán, S. (2015). fdwi evaluation of hypothalamic appetite regulation pathways in mice genetically deficient in leptin or neuropeptide y. *Neurochemical research*, 40(12):2628–2638.

[Lloyd, 1982] Lloyd, S. (1982). Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137.

[Lo et al., 1997] Lo, E. H., Pierce, A. R., Mandeville, J. B., and Rosen, B. R. (1997). Neuroprotection with nbqx in rat focal cerebral ischemia. *Stroke*, 28(2):439–447.

[Logothetis and Pfeuffer, 2004] Logothetis, N. K. and Pfeuffer, J. (2004). On the nature of the bold fmri contrast mechanism. *Magnetic resonance imaging*, 22(10):1517–1531.

[Lu et al., 2013] Lu, C.-F., Guo, W.-Y., Chang, F.-C., Huang, S.-R., Chou, Y.-C., and Wu, Y.-T. (2013). Hemodynamic segmentation of brain perfusion images with delay and dispersion effects using an expectation-maximization algorithm. *PloS one*, 8(7):e68986.

[Lu et al., 2008] Lu, C.-F., Wang, P.-S., Chou, Y.-C., Li, H.-C., Soong, B.-W., and Wu, Y.-T. (2008). Segmentation of diffusion-weighted brain images using expectation maximization algorithm initialized by hierarchical clustering. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 5502–5505. IEEE.

[Lu et al., 2017] Lu, W., Li, Z., and Chu, J. (2017). A novel computer-aided diagnosis system for breast mri based on feature selection and ensemble learning. *Computers in Biology and Medicine*, 83:157–165.

[Lungu et al., 2016] Lungu, A., Swift, A. J., Capener, D., Kiely, D., Hose, R., and Wild, J. M. (2016). Diagnosis of pulmonary hypertension from magnetic resonance imaging–based computational models and decision tree analysis. *Pulmonary circulation*, 6(2):181–190.

[Ma and Song, 2013] Ma, L. and Song, Z. J. (2013). Differentiation between low-grade and high-grade glioma using combined diffusion tensor imaging metrics. *Clinical neurology and neurosurgery*, 115(12):2489–2495.

[Madabhushi and Lee, 2016] Madabhushi, A. and Lee, G. (2016). Image analysis and machine learning in digital pathology: Challenges and opportunities. *Medical Image Analysis*, 33:170 – 175. 20th anniversary of the Medical Image Analysis journal (MedIA).

[Madabhushi and Udupa, 2006] Madabhushi, A. and Udupa, J. K. (2006). New methods of mr image intensity standardization via generalized scale. *Medical Physics*, 33(9):3426–3434.

[Mahankali et al., 2000] Mahankali, S., Liu, Y., Pu, Y., Wang, J., Chen, C.-W., Fox, P. T., and Gao, J.-H. (2000). In vivo fmri demonstration of hypothalamic function following intraperitoneal glucose administration in a rat model. *Magnetic resonance in medicine*, 43(1):155–159.

[Mahmood et al., 2015] Mahmood, Q., Chodorowski, A., Mehnert, A., Gellermann, J., and Persson, M. (2015). Unsupervised segmentation of head tissues from multi-modal mr images for eeg source localization. *Journal of digital imaging*, 28(4):499–514.

[Maintz and Viergever, 1998] Maintz, J. A. and Viergever, M. A. (1998). A survey of medical image registration. *Medical image analysis*, 2(1):1–36.

[Mandelkow et al., 2016] Mandelkow, H., de Zwart, J. A., and Duyn, J. H. (2016). Linear discriminant analysis achieves high classification accuracy for the bold fmri response to naturalistic movie stimuli. *Frontiers in human neuroscience*, 10.

[Mansfield, 1984] Mansfield, P. (1984). Real-time echo-planar imaging by nmr. *British medical bulletin*, 40(2):187.

[Mansfield and Maudsley, 1977] Mansfield, P. and Maudsley, A. A. (1977). Medical imaging by nmr. *The British journal of radiology*, 50(591):188–194.

[Marrelec et al., 2015] Marrelec, G., Messé, A., and Bellec, P. (2015). A bayesian alternative to mutual information for the hierarchical clustering of dependent random variables. *PloS one*, 10(9):e0137278.

[Matsuda et al., 1999] Matsuda, M., Liu, Y., Mahankali, S., Pu, Y., Mahankali, A., Wang, J., DeFronzo, R. A., Fox, P. T., and Gao, J.-H. (1999). Altered hypothalamic function in response to glucose ingestion in obese humans. *Diabetes*, 48(9):1801–1806.

[Matwiyoff and Brooks, 1999] Matwiyoff, N. and Brooks, W. (1999). Instrumentation. In *Magnetic resonance imaging*, pages 15–32. Mosby: St. Louis, Missouri, third edition edition.

[Mayer et al., 2010] Mayer, A., Ling, J., Mannell, M., Gasparovic, C., Phillips, J., Doezema, D., Reichard, R., and Yeo, R. (2010). A prospective diffusion tensor imaging study in mild traumatic brain injury. *Neurology*, 74(8):643–650.

[McKinley et al., 2000] McKinley, B. P., Michalek, A. M., Fenstermaker, R. A., and Plunkett, R. J. (2000). The impact of age and gender on the incidence of glial tumors in new york state from 1976–1995. *Journal of neurosurgery*, 93(6):932–939.

[McLaren et al., 2009] McLaren, C. E., Chen, W.-P., Nie, K., and Su, M.-Y. (2009). Prediction of malignant breast lesions from mri features: a comparison of artificial neural network and logistic regression techniques. *Academic radiology*, 16(7):842–851.

[McRobbie et al., 2003] McRobbie, D. W., Moore, E. A., Graves, M. J., and Prince, M. R. (2003). Mri: from picture to proton, 2003. *New York: Cambridge University Press. ISBN*, 10:0–521.

[Merisaari et al., 2009] Merisaari, H., Parkkola, R., Alhoniemi, E., Teräs, M., Lehtonen, L., Haataja, L., Lapinleimu, H., and Nevalainen, O. S. (2009). Gaussian mixture model-based segmentation of mr images taken from premature infant brains. *Journal of neuroscience methods*, 182(1):110–122.

[Mitchell, 1997] Mitchell, T. M. (1997). Machine learning. 1997. *Burr Ridge, IL: McGraw Hill*, 45(37):870–877.

[Moheet et al., 2015] Moheet, A., Mangia, S., and Seaquist, E. R. (2015). Impact of diabetes on cognitive function and brain structure. *Annals of the New York Academy of Sciences*, 1353(1):60–71.

[Moreno-Dominguez et al., 2014] Moreno-Dominguez, D., Anwander, A., and Knösche, T. R. (2014). A hierarchical method for whole-brain connectivity-based parcellation. *Human brain mapping*, 35(10):5000–5025.

[Morton et al., 2006] Morton, G., Cummings, D., Baskin, D., Barsh, G., and Schwartz, M. (2006). Central nervous system control of food intake and body weight. *Nature*, 443(7109):289–295.

[Murtagh and Contreras, 2012] Murtagh, F. and Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1):86–97.

[Murugavel and Sullivan, 2009] Murugavel, M. and Sullivan, J. M. (2009). Automatic cropping of mri rat brain volumes using pulse coupled neural networks. *Neuroimage*, 45(3):845–854.

[Naik and Patel, 2014] Naik, J. and Patel, S. (2014). Tumor detection and classification using decision tree in brain mri. *International Journal of Computer Science and Network Security (IJCSNS)*, 14(6):87.

[Nair and Halahakoon, 2016] Nair, A. and Halahakoon, C. (2016). Presentation MfD Intro and Overview. http://www.fil.ion.ucl.ac.uk/mfd_archive/2016/Schedule/Schedule.html. Accessed 2017-05-10.

[Nekooeimehr et al., 2016] Nekooeimehr, I., Lai-Yuen, S., Bao, P., Weitzenfeld, A., and Hart, S. (2016). Automated tracking, segmentation and trajectory classification of pelvic organs on dynamic mri. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, pages 2403–2406. IEEE.

[Neyns et al., 2005] Neyns, B., Sadones, J., Chaskis, C., De Ridder, M., Keyaerts, M., Veld, P., and Michotte, A. (2005). The role of chemotherapy in the treatment of low-grade glioma. a review of the literature. *Acta neurologica belgica*, 105(3):137.

[Nguyen et al., 2015] Nguyen, H. T., Jia, G., Shah, Z. K., Pohar, K., Mortazavi, A., Zynger, D. L., Wei, L., Yang, X., Clark, D., and Knopp, M. V. (2015). Prediction of chemotherapeutic response in bladder cancer using k-means clustering of dynamic contrast-enhanced (dce)-mri pharmacokinetic parameters. *Journal of Magnetic Resonance Imaging*, 41(5):1374–1382.

[Nicholas et al., 2005] Nicholas, M. K., Lukas, R., and Bangalore, S. (2005). Chemotherapy for malignant glioma. *Expert review of neurotherapeutics*, 5(sup1):41–49.

[Nichols and Holmes, 2002] Nichols, T. E. and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, 15(1):1–25.

[Nie et al., 2016] Nie, D., Wang, L., Gao, Y., and Sken, D. (2016). Fully convolutional networks for multi-modality isointense infant brain image segmentation. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, pages 1342–1345. IEEE.

[NITRC, 2007] NITRC (2007). NITRC Image Repository (NITRC-IR). http://www.nitrc.org/ir/. Accessed 2017-05-10.

[Norden et al., 2009] Norden, A. D., Drappatz, J., and Wen, P. Y. (2009). Antiangiogenic therapies for high-grade glioma. *Nature reviews. Neurology*, 5(11):610.

[Norris, 2001] Norris, D. G. (2001). The effects of microscopic tissue parameters on the diffusion weighted magnetic resonance imaging experiment. *NMR in Biomedicine*, 14(2):77–93.

[Nyúl et al., 1999] Nyúl, L. G., Udupa, J. K., et al. (1999). On standardizing the mr image intensity scale. *Magnetic Resonance in Medicine*, 42:1072–1081.

[Nyúl et al., 2000] Nyúl, L. G., Udupa, J. K., and Zhang, X. (2000). New variants of a method of mri scale standardization. *IEEE transactions on medical imaging*, 19(2):143–150.

[Oguz et al., 2014] Oguz, I., Zhang, H., Rumple, A., and Sonka, M. (2014). Rats: rapid automatic tissue segmentation in rodent brain mri. *Journal of neuroscience methods*, 221:175–182.

[Oikonomou and Blekas, 2013] Oikonomou, V. P. and Blekas, K. (2013). An adaptive regression mixture model for fmri cluster analysis. *IEEE transactions on medical imaging*, 32(4):649–659.

[Oliveira and Tavares, 2014] Oliveira, F. P. and Tavares, J. M. R. (2014). Medical image registration: a review. *Computer methods in biomechanics and biomedical engineering*, 17(2):73–93.

[Ortiz et al., 2013] Ortiz, A., Gorriz, J. M., Ramírez, J., Salas-Gonzalez, D., Initiative, A. D. N., et al. (2013). Improving mri segmentation with probabilistic ghsom and multiobjective optimization. *Neurocomputing*, 114:118–131.

[Park et al., 2017] Park, C.-A., Kang, C.-K., Kim, Y.-B., and Cho, Z.-H. (2017). Advances in mr angiography with 7t mri: From microvascular imaging to functional angiography. *NeuroImage*.

[Parkinson et al., 2009] Parkinson, J. R., Chaudhri, O. B., and Bell, J. D. (2009). Imaging appetite-regulating pathways in the central nervous system using manganese-enhanced magnetic resonance imaging. *Neuroendocrinology*, 89(2):121–130.

[Patel et al., 2014] Patel, A. P., Tirosh, I., Trombetta, J. J., Shalek, A. K., Gillespie, S. M., Wakimoto, H., Cahill, D. P., Nahed, B. V., Curry, W. T., Martuza, R. L., et al. (2014). Single-cell rna-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science*, 344(6190):1396–1401.

[Patel et al., 2016] Patel, M. J., Khalaf, A., and Aizenstein, H. J. (2016). Studying depression using imaging and machine learning methods. *NeuroImage: Clinical*, 10:115–123.

[Pathak, 2009] Pathak, A. P. (2009). Magnetic resonance susceptibility based perfusion imaging of tumors using iron oxide nanoparticles. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, 1(1):84–97.

[Paxinos and Franklin, 2001] Paxinos, G. and Franklin, K. B. (2001). *The mouse brain in stereotaxic coordinates*. New York, NY.

[Paxinos and Watson, 2004] Paxinos, G. and Watson, C. (2004). *The rat brain in stereotaxic coordinates*. New York, NY.

[Peng et al., 2005] Peng, H., Long, F., and Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8):1226–1238.

[Penny et al., 2011] Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., and Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Academic press.

[Pereira et al., 2016] Pereira, S., Pinto, A., Alves, V., and Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in mri images. *IEEE transactions on medical imaging*, 35(5):1240–1251.

[Pérez and Sánchez-Montañés, 2007] Pérez, O. and Sánchez-Montañés, M. (2007). A new learning strategy for classification problems with different training and test

distributions. In *Proceedings of the 9th International Work Conference on Artificial Neural Networks*, IWANN'07, pages 178–185, Berlin, Heidelberg. Springer-Verlag.

[Pérez and Sánchez-Montañés, 2009] Pérez, Ó. and Sánchez-Montañés, M. (2009). Class prediction in test sets with shifted distributions. In *Encyclopedia of Artificial Intelligence*, pages 282–288. IGI Global.

[Pérez-Carro et al., 2014] Pérez-Carro, R., Cauli, O., and López-Larrubia, P. (2014). Multiparametric magnetic resonance in the assessment of the gender differences in a high-grade glioma rat model. *EJNMMI research*, 4(1):1.

[Petersson et al., 1999] Petersson, K. M., Nichols, T. E., Poline, J.-B., and Holmes, A. P. (1999). Statistical limitations in functional neuroimaging. i. non-inferential methods and statistical models. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 354(1387):1239–1260.

[Plant et al., 2010] Plant, C., Teipel, S. J., Oswald, A., Böhm, C., Meindl, T., Mourao-Miranda, J., Bokde, A. W., Hampel, H., and Ewers, M. (2010). Automated detection of brain atrophy patterns based on mri for the prediction of alzheimer's disease. *Neuroimage*, 50(1):162–174.

[Purcell et al., 1946] Purcell, E. M., Torrey, H., and Pound, R. V. (1946). Resonance absorption by nuclear magnetic moments in a solid. *Physical review*, 69(1-2):37.

[Puttick et al., 2015] Puttick, S., Bell, C., Dowson, N., Rose, S., and Fay, M. (2015). Pet, mri, and simultaneous pet/mri in the development of diagnostic and therapeutic strategies for glioma. *Drug discovery today*, 20(3):306–317.

[Quinlan, 1986] Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1):81–106.

[Raftery, 1986] Raftery, A. E. (1986). A note on bayes factors for log-linear contingency table models with vague prior information. *Journal of the Royal Statistical Society, Series B*, 48:249–250.

[Raizer et al., 2015] Raizer, J., Parsa, A., et al. (2015). *Current understanding and treatment of gliomas*. Springer.

[Ranjith et al., 2015] Ranjith, G., Parvathy, R., Vikas, V., Chandrasekharan, K., and Nair, S. (2015). Machine learning methods for the classification of gliomas: Initial results using features extracted from mr spectroscopy. *The neuroradiology journal*, page 1971400915576637.

[Rao et al., 2011] Rao, A., Lee, Y., Gass, A., and Monsch, A. (2011). Classification of alzheimer's disease from structural mri using sparse logistic regression with optional spatial regularization. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 4499–4502. IEEE.

[Razlighi et al., 2013] Razlighi, Q. R., Kehtarnavaz, N., and Yousefi, S. (2013). Evaluating similarity measures for brain image registration. *Journal of visual communication and image representation*, 24(7):977–987.

[Ren et al., 2016] Ren, M., Liu, P., Wang, Z., and Yi, J. (2016). A self-adaptive fuzzy c-means algorithm for determining the optimal number of clusters. *Computational Intelligence and Neuroscience*, 2016.

[Rendón et al., 2011] Rendón, E., Abundez, I., Arizmendi, A., and Quiroz, E. M. (2011). Internal versus external cluster validation indexes. *International Journal of computers and communications*, 5(1):27–34.

[Rousseeuw, 1987] Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65.

[Roweis and Saul, 2000] Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500):2323–2326.

[Roy and Bhattacharyya, 2015] Roy, S. and Bhattacharyya, D. K. (2015). Segmentation of cortical gray and white matters from mri using density based clustering approach. In *Advanced Computing and Communication (ISACC), 2015 International Symposium on*, pages 181–185. IEEE.

[Rueckert et al., 2016] Rueckert, D., Glocker, B., and Kainz, B. (2016). Learning clinically useful information from images: Past, present and future. *Medical Image Analysis*, 33:13 – 18. 20th anniversary of the Medical Image Analysis journal (MedIA).

[Ryali et al., 2015] Ryali, S., Chen, T., Padmanabhan, A., Cai, W., and Menon, V. (2015). Development and validation of consensus clustering-based framework for brain segmentation using resting fmri. *Journal of neuroscience methods*, 240:128–140.

[Ryken et al., 2015] Ryken, T. C., Parney, I., Buatti, J., Kalkanis, S. N., and Olson, J. J. (2015). The role of radiotherapy in the management of patients with diffuse low grade glioma. *Journal of neuro-oncology*, 125(3):551–583.

[Saeys et al., 2007] Saeys, Y., Inza, I., and Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *bioinformatics*, 23(19):2507–2517.

[Sánchez-Montañés and Baroni, 2010] Sánchez-Montañés, M. and Baroni, F. (2010). Invariants and specificities of neuronal activity revealed by a novel clustering algorithm for multiple data sets. In *FENS Abstr.*, volume 5 of *FENS Forum'10, Abstract number 059.30*. Federation of European Neurosciences Societies (FENS).

[Sander et al., 1998] Sander, J., Ester, M., Kriegel, H.-P., and Xu, X. (1998). Density-based clustering in spatial databases: The algorithm gdbscan and its applications. *Data mining and knowledge discovery*, 2(2):169–194.

[Sawiak et al., 2009] Sawiak, S., Wood, N., Williams, G., Morton, A., and Carpenter, T. (2009). Spmmouse: A new toolbox for spm in the animal brain. In *ISMRM 17th Scientific Meeting & Exhibition, April*, pages 18–24.

[Schäfer et al., 2013] Schäfer, M.-L., Maurer, M. H., Synowitz, M., Wüstefeld, J., Marnitz, T., Streitparth, F., and Wiener, E. (2013). Low-grade (who ii) and anaplastic (who iii) gliomas: differences in morphology and mri signal intensities. *European radiology*, 23(10):2846–2853.

[Schapire, 1990] Schapire, R. E. (1990). The strength of weak learnability. *Machine learning*, 5(2):197–227.

[Schenck, 1996] Schenck, J. F. (1996). The role of magnetic susceptibility in magnetic resonance imaging: Mri magnetic compatibility of the first and second kinds. *Medical physics*, 23(6):815–850.

[Schmidhuber, 2015] Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61:85–117.

[Schwartz and Porte, 2005] Schwartz, M. W. and Porte, D. (2005). Diabetes, obesity, and the brain. *Science*, 307(5708):375–379.

[Schwartz et al., 2000] Schwartz, M. W., Woods, S. C., Porte, D., Seeley, R. J., and Baskin, D. G. (2000). Central nervous system control of food intake. *Nature*, 404(6778):661–671.

[Ségonne et al., 2004] Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., and Fischl, B. (2004). A hybrid approach to the skull stripping problem in mri. *Neuroimage*, 22(3):1060–1075.

[Shannon, 1948] Shannon, C. (1948). A mathematical theory of communication, bell system technical journal 27: 379-423 and 623–656. *Mathematical Reviews (MathSciNet): MR10, 133e*.

[Shattuck et al., 2009] Shattuck, D. W., Prasad, G., Mirza, M., Narr, K. L., and Toga, A. W. (2009). Online resource for validation of brain segmentation methods. *NeuroImage*, 45(2):431–439.

[Shattuck et al., 2001] Shattuck, D. W., Sandor-Leahy, S. R., Schaper, K. A., Rottenberg, D. A., and Leahy, R. M. (2001). Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13(5):856–876.

[Shen et al., 2017] Shen, D., Wu, G., and Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19(1).

[Shen et al., 2011] Shen, T.-W., Hsin, Y.-L., and Harnod, T. (2011). Using fuzzy c-means index matrix to depict for the focal cortical dysplasia region on t1 brain mri images. In *Medical Information & Communication Technology (ISMICT), 2011 5th International Symposium on*, pages 117–121. IEEE.

[Simon et al., 2012] Simon, D., Fritzsche, K. H., Thieke, C., Klein, J., Parzer, P., Weber, M.-A., and Stieltjes, B. (2012). Diffusion-weighted imaging-based probabilistic segmentation of high-and low-proliferative areas in high-grade gliomas. *Cancer Imaging*, 12(1):89–99.

[Simpson et al., 2013] Simpson, I. J., Woolrich, M. W., Andersson, J. R., Groves, A. R., and Schnabel, J. (2013). Ensemble learning incorporating uncertain registration. *IEEE transactions on medical imaging*, 32(4):748–756.

[Singh et al., 2004] Singh, S. K., Hawkins, C., Clarke, I. D., Squire, J. A., Bayani, J., Hide, T., Henkelman, R. M., Cusimano, M. D., and Dirks, P. B. (2004). Identification of human brain tumour initiating cells. *nature*, 432(7015):396–401.

[Sled et al., 1998] Sled, J. G., Zijdenbos, A. P., and Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in mri data. *IEEE transactions on medical imaging*, 17(1):87–97.

[Slomka and Baum, 2009] Slomka, P. J. and Baum, R. P. (2009). Multimodality image registration with software: state-of-the-art. *European journal of nuclear medicine and molecular imaging*, 36(1):44.

[Smith, 2002] Smith, S. M. (2002). Fast robust automated brain extraction. *Human brain mapping*, 17(3):143–155.

[Somasundaram and Kalavathi, 2011] Somasundaram, K. and Kalavathi, P. (2011). Medical image contrast enhancement based on gamma correction. *Int J Knowl Manag e-learning*, 3(1):15–18.

[Somasundaram and Kalavathi, 2012] Somasundaram, K. and Kalavathi, P. (2012). Analysis of imaging artifacts in mr brain images. *Oriental J Comput Sci Technol*, 5(1):135–141.

[Sorensen et al., 2008] Sorensen, A. G., Batchelor, T. T., Wen, P. Y., Zhang, W.-T., and Jain, R. K. (2008). Response criteria for glioma. *Nature Clinical Practice Oncology*, 5(11):634–644.

[Sotak, 2002] Sotak, C. H. (2002). The role of diffusion tensor imaging in the evaluation of ischemic brain injury–a review. *NMR in Biomedicine*, 15(7-8):561–569.

[Stejskal, 1965] Stejskal, E. (1965). Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow. *The Journal of Chemical Physics*, 43(10):3597–3603.

[Steven et al., 2014] Steven, A. J., Zhuo, J., and Melhem, E. R. (2014). Diffusion kurtosis imaging: an emerging technique for evaluating the microstructural environment of the brain. *American Journal of Roentgenology*, 202(1):W26–W33.

[Strehl and Ghosh, 2002] Strehl, A. and Ghosh, J. (2002). Cluster ensembles—a knowledge reuse framework for combining multiple partitions. *Journal of machine learning research*, 3(Dec):583–617.

[Strother, 2006] Strother, S. C. (2006). Evaluating fmri preprocessing pipelines. *IEEE Engineering in Medicine and Biology Magazine*, 25(2):27–41.

[Strumia et al., 2012] Strumia, M., Ramantani, G., Mader, I., Henning, J., Bai, L., and Hadjidemetriou, S. (2012). Analysis of structural mri data for the localisation of focal cortical dysplasia in epilepsy. In *Workshop on Clinical Image-Based Procedures*, pages 25–32. Springer.

[Stylli et al., 2015] Stylli, S. S., Luwor, R. B., Ware, T. M., Tan, F., and Kaye, A. H. (2015). Mouse models of glioma. *Journal of Clinical Neuroscience*, 22(4):619–626.

[Suk et al., 2017] Suk, H.-I., Lee, S.-W., Shen, D., Initiative, A. D. N., et al. (2017). Deep ensemble learning of sparse regression models for brain disease diagnosis. *Medical image analysis*, 37:101–113.

[Sun et al., 2012] Sun, T., Warrington, N. M., and Rubin, J. B. (2012). Why does jack, and not jill, break his crown? sex disparity in brain tumors. *Biology of sex differences*, 3(1):3.

[Svolos et al., 2013] Svolos, P., Tsolaki, E., Theodorou, K., Fountas, K., Kapsalaki, E., Fezoulidis, I., and Tsougos, I. (2013). Classification methods for the differentiation of atypical meningiomas using diffusion and perfusion techniques at 3-t mri. *Clinical imaging*, 37(5):856–864.

[Sweeney et al., 2014] Sweeney, E. M., Vogelstein, J. T., Cuzzocreo, J. L., Calabresi, P. A., Reich, D. S., Crainiceanu, C. M., and Shinohara, R. T. (2014). A comparison of supervised machine learning algorithms and feature vectors for ms lesion segmentation using multimodal structural mri. *PloS one*, 9(4):e95753.

[Sweeney et al., 2015] Sweeney, T. E., Chen, A. C., and Gevaert, O. (2015). Combined mapping of multiple clustering algorithms (communal): A robust method for selection of cluster number, k. *Scientific reports*, 5.

[Syková and Nicholson, 2008] Syková, E. and Nicholson, C. (2008). Diffusion in brain extracellular space. *Physiological reviews*, 88(4):1277–1340.

[Szafer et al., 1995] Szafer, A., Zhong, J., and Gore, J. C. (1995). Theoretical model for water diffusion in tissues. *Magnetic resonance in medicine*, 33(5):697–712.

[Taal et al., 2015] Taal, W., Bromberg, J. E., and van den Bent, M. J. (2015). Chemotherapy in glioma. *CNS oncology*, 4(3):179–192.

[Talairach and Tournoux, 1988] Talairach, J. and Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: An approach to cerebral imaging. *Thieme Medical*.

[Tanoori et al., 2011] Tanoori, B., Azimifar, Z., Shakibafar, A., and Katebi, S. (2011). Brain volumetry: an active contour model-based segmentation followed by svm-based classification. *Computers in biology and medicine*, 41(8):619–632.

[Tataranni et al., 1999] Tataranni, P. A., Gautier, J.-F., Chen, K., Uecker, A., Bandy, D., Salbe, A. D., Pratley, R. E., Lawson, M., Reiman, E. M., and Ravussin, E. (1999). Neuroanatomical correlates of hunger and satiation in humans using positron emission tomography. *Proceedings of the National Academy of Sciences*, 96(8):4569–4574.

[Tellaroli et al., 2016] Tellaroli, P., Bazzi, M., Donato, M., Brazzale, A. R., and Drăghici, S. (2016). Cross-clustering: A partial clustering algorithm with automatic estimation of the number of clusters. *PloS one*, 11(3):e0152333.

[Tenenbaum et al., 2000] Tenenbaum, J. B., De Silva, V., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323.

[Therasse et al., 2000] Therasse, P., Arbuck, S. G., Eisenhauer, E. A., Wanders, J., Kaplan, R. S., Rubinstein, L., Verweij, J., Van Glabbeke, M., van Oosterom, A. T., Christian, M. C., et al. (2000). New guidelines to evaluate the response to treatment in solid tumors. *Journal of the National Cancer Institute*, 92(3):205–216.

[Tibshirani, 1996] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288.

[Triadyaksa et al., 2016] Triadyaksa, P., Prakken, N. H., Overbosch, J., Peters, R. B., van Swieten, J. M., Oudkerk, M., and Sijens, P. E. (2016). Semi-automated myocardial segmentation of bright blood multi-gradient echo images improves reproducibility of myocardial contours and t2* determination. *Magnetic Resonance Materials in Physics, Biology and Medicine*, pages 1–16.

[Tuch, 2004] Tuch, D. S. (2004). Q-ball imaging. *Magnetic resonance in medicine*, 52(6):1358–1372.

[Tustison et al., 2010] Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., and Gee, J. C. (2010). N4itk: improved n3 bias correction. *IEEE transactions on medical imaging*, 29(6):1310–1320.

[Van Buchem and Tofts, 2000] Van Buchem, M. and Tofts, P. S. (2000). Magnetization transfer imaging. *Neuroimaging Clinics of North America*, 10(4):771–88.

[Vergara et al., 2017] Vergara, V. M., Mayer, A. R., Damaraju, E., Hutchison, K., and Calhoun, V. D. (2017). The effect of preprocessing pipelines in subject classification and detection of abnormal resting state functional network connectivity using group ica. *Neuroimage*, 145:365–376.

[Verhaak, 2016] Verhaak, R. (2016). Moving the needle: Optimizing classification for glioma. *Science translational medicine*, 8(350):350fs14.

[Victoria et al., 2015] Victoria, Y. Y., Nguyen, D., Pajonk, F., Kupelian, P., Kaprealian, T., Selch, M., Low, D. A., and Sheng, K. (2015). Incorporating cancer stem cells in radiation therapy treatment response modeling and the implication in glioblastoma multiforme treatment resistance. *International Journal of Radiation Oncology* Biology* Physics*, 91(4):866–875.

[Viergever et al., 2016] Viergever, M. A., Maintz, J. A., Klein, S., Murphy, K., Staring, M., and Pluim, J. P. (2016). A survey of medical image registration–under review. *Medical image analysis*, 33:140–144.

[Vovk et al., 2007] Vovk, U., Pernus, F., and Likar, B. (2007). A review of methods for correction of intensity inhomogeneity in mri. *IEEE transactions on medical imaging*, 26(3):405–421.

[Vrooman et al., 2007] Vrooman, H. A., Cocosco, C. A., van der Lijn, F., Stokking, R., Ikram, M. A., Vernooij, M. W., Breteler, M. M., and Niessen, W. J. (2007). Multi-spectral brain tissue segmentation using automatically trained k-nearest-neighbor classification. *Neuroimage*, 37(1):71–81.

[Wachinger et al., 2017] Wachinger, C., Reuter, M., and Klein, T. (2017). Deepnat: Deep convolutional neural network for segmenting neuroanatomy. *NeuroImage*.

[Wang et al., 2016] Wang, L., Chitiboi, T., Meine, H., Günther, M., and Hahn, H. K. (2016). Principles and methods for automatic and semi-automatic tissue segmentation in mri data. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 29(2):95–110.

[Wang and Summers, 2012] Wang, S. and Summers, R. M. (2012). Machine learning and radiology. *Medical image analysis*, 16(5):933–951.

[Wang and Li, 2013] Wang, Y. and Li, T.-Q. (2013). Analysis of whole-brain resting-state fmri data using hierarchical clustering approach. *PLoS One*, 8(10):e76315.

[Ward and Thompson, 2012] Ward, P. S. and Thompson, C. B. (2012). Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. *Cancer cell*, 21(3):297–308.

[Way et al., 2010] Way, T. W., Sahiner, B., Hadjiiski, L. M., and Chan, H.-P. (2010). Effect of finite sample size on feature selection and classification: a simulation study. *Medical physics*, 37(2):907–920.

[Weinert, 1979] Weinert, H. (1979). Bryson, ae/ho, y.-c., applied optimal control, optimization, estimation, and control. new york-london-sydney-toronto. john wiley & sons. 1975. 481 s.,£ 10.90. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 59(8):402–402.

[Williams and Elmquist, 2012] Williams, K. W. and Elmquist, J. K. (2012). From neuroanatomy to behavior: central integration of peripheral signals regulating feeding behavior. *Nature neuroscience*, 15(10):1350–1355.

[Woitek et al., 2017] Woitek, R., Spick, C., Schernthaner, M., Rudas, M., Kapetas, P., Bernathova, M., Furtner, J., Pinker, K., Helbich, T. H., and Baltzer, P. A. (2017). A simple classification system (the tree flowchart) for breast mri can reduce the number of unnecessary biopsies in mri-only lesions. *European Radiology*, pages 1–11.

[Wolff and Balaban, 1989] Wolff, S. D. and Balaban, R. S. (1989). Magnetization transfer contrast (mtc) and tissue water proton relaxation in vivo. *Magnetic resonance in medicine*, 10(1):135–144.

[Wood and Wehrli, 1999] Wood, M. and Wehrli, F. (1999). Principles of magnetic resonance imaging. In *Magnetic resonance imaging*, pages 1–14. Mosby: St. Louis, Missouri, third edition edition.

[Worsley et al., 1992] Worsley, K. J., Evans, A. C., Marrett, S., and Neelin, P. (1992). A three-dimensional statistical analysis for cbf activation studies in human brain. *Journal of Cerebral Blood Flow & Metabolism*, 12(6):900–918.

[Wu et al., 2015a] Wu, J., Qian, Z., Tao, L., Yin, J., Ding, S., Zhang, Y., and Yu, Z. (2015a). Resting state fmri feature-based cerebral glioma grading by support vector machine. *International journal of computer assisted radiology and surgery*, 10(7):1167–1174.

[Wu et al., 2015b] Wu, Z., Matsui, O., Kitao, A., Kozaka, K., Koda, W., Kobayashi, S., Ryu, Y., Minami, T., Sanada, J., and Gabata, T. (2015b). Hepatitis c related chronic liver cirrhosis: feasibility of texture analysis of mr images for classification of fibrosis stage and necroinflammatory activity grade. *PloS one*, 10(3):e0118297.

[Xia et al., 2015] Xia, M., Song, S., Yao, L., and Long, Z. (2015). An empirical comparison of different lda methods in fmri-based brain states decoding. *Bio-Medical Materials and Engineering*, 26(s1):S1185–S1192.

[Xu and Wunsch, 2005] Xu, R. and Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on neural networks*, 16(3):645–678.

[Yablonskiy and Sukstanskii, 2010] Yablonskiy, D. A. and Sukstanskii, A. L. (2010). Theoretical models of the diffusion weighted mr signal. *NMR in Biomedicine*, 23(7):661–681.

[Yang et al., 2011] Yang, C.-A., Kaveh, M., and Erickson, B. J. (2011). Automated detection of focal cortical dysplasia lesions on t1-weighted mri using volume-based distributional features. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pages 865–870. IEEE.

[Yiğit et al., 2017] Yiğit, H., Önder, A., Özgür, S., Aycan, Z., Karademir, S., and Doğan, V. (2017). Cardiac mri and 3d contrast-enhanced mr angiography in pediatric and young adult patients with turner syndrome. *Turkish journal of medical sciences*, 47(1):127–133.

[Yun et al., 2012] Yun, J., Mackenzie, M., Rathee, S., Robinson, D., and Fallone, B. (2012). An artificial neural network (ann)-based lung-tumor motion predictor for intrafractional mr tumor tracking. *Medical physics*, 39(7):4423–4433.

[Zhang and Ma, 2012] Zhang, C. and Ma, Y. (2012). *Ensemble machine learning*, volume 1. Springer.

[Zhang et al., 2015] Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., and Shen, D. (2015). Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *NeuroImage*, 108:214–224.

[Zheng et al., 2009] Zheng, W., Chee, M. W., and Zagorodnov, V. (2009). Improvement of brain segmentation accuracy by optimizing non-uniformity correction using n3. *NeuroImage*, 48(1):73–83.

[Zhou et al., 2015] Zhou, X., Wang, S., Xu, W., Ji, G., Phillips, P., Sun, P., and Zhang, Y. (2015). Detection of pathological brain in mri scanning based on wavelet-entropy and naive bayes classifier. In *International Conference on Bioinformatics and Biomedical Engineering*, pages 201–209. Springer.

[Zöllner et al., 2012] Zöllner, F. G., Emblem, K. E., and Schad, L. R. (2012). Svm-based glioma grading: optimization by feature reduction analysis. *Zeitschrift für medizinische Physik*, 22(3):205–214.

[Zou and Hastie, 2005] Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.