Doctoral
Thesis

Enhancing brain tumor diagnosis with Synthetic MRI

Elisa Moya Sáez

Advisors:
Prof. Carlos Alberola López
Prof. Rodrigo de Luis García

Universidad de Valladolid

Escuela de Doctorado **Universidad de Valladolid**

Programa de Doctorado en Tecnologías de la Información y las Telecomunicaciones

Tesis Doctoral

Enhancing brain tumor diagnosis with Synthetic MRI

Presentada por:

Elisa Moya Sáez

para optar al grado de Doctora por la Universidad de Valladolid

Dirigida por:
Dr. Carlos Alberola López
Dr. Rodrigo de Luis García

Valladolid, España, 2024

Título:
Title:
Enhancing brain tumor diagnosis with Synthetic MRI

Autor:
Author:
Msc. Elisa Moya Sáez

Director:
Director:
Dr. Carlos Alberola López

Codirector:
Codirector:
Dr. Rodrigo de Luis García

Departamento:
Department:
Teoría de la Señal Comunicaciones e Ingeniería Telemática

Información de contacto:
Contact information:

🖂 **Elisa Moya Sáez**
Laboratorio de Procesado de Imagen (LPI)
Dept. Teoría de la Señal Comunicaciones e Ingeniería Telemática
ETS. Ingenieros de Telecomunicación
Universidad de Valladolid
Campus Miguel Delibes sn. 47011
Valladolid, Spain
🖃 `elisa.moya@uva.es`
🌍 `http://www.lpi.tel.uva.es/emoysae`

“Las cosas podían haber sucedido de cualquier otra manera y, sin embargo, sucedieron así.”

Miguel Delibes, *El Camino*

Summary

Malignant gliomas are the most common primary brain tumors in adults. This family of brain tumors includes different types that differ in their genetic characteristics and prognostic outcomes, the latter being generally unfavorable. Survival is especially poor in high-grade gliomas such as glioblastomas, so in those cases predicting the expected survival is crucial for efficient surgery and treatment planning.

As for diagnosis, in clinical practice this is commonly performed by magnetic resonance imaging (MRI) and, in particular, by visual inspection of the weighted images. The standard MRI protocol for brain tumor assessment includes (at least) four different weighted images: T1-weighted, T2-weighted, FLAIR and T1-weighted after injection of a gadolinium-based contrast agent (GBCA). The latter is used to evaluate blood brain barrier breakdown, a condition displayed on the image as signal enhancement caused by the contrast agent extravasation into the perivascular space. This diagnostic procedure has two main limitations; on the one hand, the qualitative nature of the weighted images hinders the usage of quantitative methods. On the other hand, the usage of GBCAs can trigger adverse effects that under certain circumstances can be severe, in addition to increasing the scan time and cost.

In contrast, quantitative MRI is based on the computation of the tissue magnetic properties themselves, collectively known as parametric maps. These properties are the longitudinal relaxation time (T1), transverse relaxation time (T2) and proton density (PD). Parametric maps present an absolute scale and are generally considered more robust than weighted images. Recently, a new paradigm, Synthetic MRI, has gained popularity; it is based on the T1, T2, and PD parametric maps computation, followed by the synthesis of several weighted images from these maps. As a result, this procedure can enhance efficiency and diagnosis.

Parametric maps computation can be performed by means of traditional relaxometry sequences. These sequences consist in the acquisition of a set of weighted images varying a particular acquisition parameter, followed by the voxelwise fitting of these images to a known relaxation model with estimation-based procedures. However, the long scan time of these sequences hampers their usage in clinical practice. Alternatively, fast multiparametric mapping techniques have been recently proposed. Although these techniques are faster than relaxometry sequences, they still imply a not negligible acquisition time. In addition, they usually require specific sequences or commercial products that are scarcely available on clinical scanners.

Deep learning approaches could also be an alternative for the computation of parametric maps from conventional and, therefore, widely available weighted images. Thus, parametric maps could be easily computed both on pre-existing databases as well as on new acquisitions without increasing scan time. However, the lack of public datasets containing weighted images and their corresponding parametric maps could be one of the main limitations that hinders the usage of deep learning.

In this Thesis we propose to enhance the diagnosis of brain tumors following a Synthetic MRI paradigm. The computation of T1, T2, and PD parametric maps have been performed with deep learning from conventional weighted images acquired with routine protocols. After that, different types of weighted images have been successfully synthesized out of the parametric maps. Several solutions have been proposed to overcome the lack of public datasets with parametric maps. In particular, training based on synthetic data and self-supervised learning strategies, which enable the training from only weighted images. In addition, both the synthesized weighted images and the computed parametric maps have been employed in different applications to improve brain tumor diagnosis. Specifically, predicting both the expected survival of glioblastoma patients and the post-contrast T1-weighted-enhanced tissues without the injection of a GBCA.

Resumen

El glioma maligno es el tumor cerebral primario más común en adultos. Esta familia de tumores cerebrales incluye distintos tipos que se diferencian por sus características genéticas y el pronóstico esperado, siendo este último en general desfavorable. La supervivencia esperada es especialmente baja en los gliomas de alto grado como los glioblastomas, por lo que en esos casos predecir la supervivencia es una tarea crucial para la planificación eficiente de cirugía y tratamiento.

En lo que se refiere al diagnóstico, en la práctica clínica este es comúnmente realizado mediante imagen de resonancia magnética (IRM) y, en particular, mediante inspección visual de imágenes ponderadas. El protocolo estándar de resonancia magnética (RM) para evaluar tumores cerebrales incluye, al menos, cuatro imágenes ponderadas distintas: ponderada en T1, ponderada en T2, FLAIR y ponderada en T1 después de la inyección de un agente de contraste a base de gadolinio. Esta última se utiliza para evaluar la ruptura de la barrera hematoencefálica, condición manifestada en la imagen como realce de señal causado por la fuga de contraste al espacio perivascular. Este procedimiento diagnóstico presenta dos limitaciones principales. Por un lado, la naturaleza cualitativa de las imágenes ponderadas dificulta el uso de métodos cuantitativos. Por otro lado, el uso de agentes de contraste puede desencadenar efectos adversos que bajo ciertas circunstancias pueden ser severos, además de incrementar la duración y el coste de las pruebas.

Por el contrario la IRM cuantitativa se basa en calcular directamente las propiedades magnéticas de los tejidos conocidas como mapas paramétricos. Estas son el tiempo de relajación longitudinal (T1), tiempo de relajación transversal (T2) y densidad protónica (PD). Estos parámetros presentan una escala absoluta y en general son considerados más robustos que las imágenes ponderadas. Recientemente, un nuevo paradigma, la IRM sintética, ha cobrado relevancia; este se basa en la computación de mapas paramétricos de T1, T2 y PD, seguido de la síntesis de imágenes ponderadas a partir de los mapas. Este procedimiento tiene la ventaja de mejorar la eficiencia de las pruebas y facilitar el diagnóstico.

El cálculo de los mapas paramétricos puede realizarse mediante secuencias clásicas de relaxometría, sin embargo, su utilidad en la práctica clínica está limitada por su larga duración. Como alternativa, recientemente se han propuesto técnicas multiparamétricas rápidas que alivian este problema aunque aún requieren un tiempo de adquisición no despreciable. Además, estás técnicas están normalmente asociadas a secuencias específicas o productos comerciales que en muchas ocasiones no están disponibles en los escáneres clínicos.

Las técnicas de aprendizaje profundo podrían ser una alternativa para el cálculo de mapas paramétricos a partir de imágenes ponderadas convencionales y, por tanto, ampliamente disponibles. De este modo, los mapas paramétricos podrían calcularse fácilmente tanto en bases de datos ya existentes, así como en nuevas adquisiciones sin incrementar el tiempo de las pruebas. Sin embargo, la falta de conjuntos de datos públicos que contengan imágenes ponderadas y los correspondientes mapas paramétricos asociados podría ser una de las principales limitaciones que dificultan el uso de estas técnicas, ya que estas requieren, en principio, grandes conjuntos de datos para su entrenamiento.

En esta Tesis Doctoral proponemos mejorar el diagnóstico de tumores cerebrales siguiendo el paradigma de IRM sintética. El cálculo de los mapas paramétricos se ha realizado a partir de imágenes ponderadas convencionales adquiridas con protocolos rutinarios haciendo uso de técnicas de aprendizaje profundo. Posteriormente, se han sintetizado con éxito diferentes tipos de imágenes ponderadas a partir de los mapas paramétricos. Distintas soluciones se han propuesto para solventar el problema de la falta de conjuntos de datos públicos con mapas paramétricos. En concreto, entrenamientos basados en datos sintéticos y estrategias de aprendizaje auto-supervisado para realizar el entrenamiento a partir únicamente de imágenes ponderadas. Además, tanto las imágenes ponderadas sintetizadas como los mapas paramétricos computados se han empleado en diferentes aplicaciones para mejorar el diagnóstico de tumores cerebrales. En concreto, se han utilizado tanto para predecir la supervivencia esperada de los pacientes con glioblastoma, como para predecir los tejidos tumorales con realce sin la utilización de agentes de contraste.

Acknowledgements

This thesis is a reflection of the collective efforts and support of many individuals, who I would like to thank. Firstly, I would like to express my gratitude to my advisors Prof. Carlos Alberola López and Prof. Rodrigo de Luis García for their unwavering support and guidance throughout the entire process. Their expertise, encouragement, and constructive feedback have been instrumental in shaping this Thesis.

I definitely need to include Prof. Juan Antonio Hernández Tamames in these acknowledgments. First, for giving me the opportunity to do my research stay at the Erasmus MC and, also, for his kindness and encouragement when I was there. His invaluable insights have deeply enriched my research experience and contributed to the overall success of this project.

My sponsor, the Asociación Española Contra el Cáncer, deserves a special mention for my predoctoral grant and, also, all the people that make that association possible: professionals, volunteers, members, and the more than forty six thousand people that every year fill the streets of Valladolid and dye them green. Also, the Universidad de Valladolid for the facilities and resources.

I also want to thank all the staff that has been at the LPI during these years who have offered me their assistance and shared their knowledge. I would like to mention Rafael Navarro for the design and development of the radiomics system employed in this Thesis. Without any doubt, the best memories of this Thesis are thanks to all of you: Dani, Iñaki, Santi, Óscar, Elena, Rosa, Susana, Patri, Tomasz, Álvaro, Emilio, Rafa, Tino, and Pablo.

A su vez, me gustaría expresar mis más sinceros agradecimientos a mi familia y amigos. A mis padres por su apoyo y amor incondicional. A mi hermana y Arturo por estar siempre ahí. A mi tío por transmitirme su calma. A Jorge y Mario, que espero algún día puedan leer esto, por alegrarme cada día. A mis abuelos y abuelas, Robus, Nino, Antonina y Esther, que tengo la certeza de que estarían orgullosos de mí. A mis tí@s y primas por su cariño y apoyo. Y por último, a Óscar, por su amor, apoyo y paciencia en los mejores y peores momentos y por sus sabios consejos que hacen que esta tesis, hoy, pueda ser una realidad. Mil gracias a todos vosotros. Os quiero.

Thank you each and every one of you,

Elisa

Acronyms

This section lists (in alphabetic order) all the acronyms that have been used in this Thesis dissertation.

ABN abnormal tissue. 87, 89, 110, 143

APT amide proton transfer. 27

ASL arterial spin labeling. 27

AUC area under the curve. 73, 74, 76, 79, 143, 144

BBB blood brain barrier. 4, 25, 85, 94, 98, 120

BTB blood tumor barrier. 25

CNN convolutional neural network. 9, 15, 34, 37, 38, 43–46, 49, 50, 52, 53, 57–59, 63, 64, 70–72, 74–76, 79, 87, 98, 99, 103, 107–109, 111, 114, 120

CNR contrast-to-noise ratio. 48, 52, 54–56, 61, 62

CNS central nervous system. 3, 25

CORR correlation coefficient. 48, 56, 57

CSF cerebrospinal fluid. 44, 47, 52–55, 61, 63

CT computerized tomography. 4, 37, 38

DFT discrete Fourier transform. 21

DL deep learning. 7–9, 15, 35, 43, 58, 69, 70, 78, 85–95, 104, 107, 108, 110, 117–122

FCNN fully convolutional neural network. 35, 37

FID free induction decay. 19

FLAIR fluid attenuated inversion recovery. 23

FT Fourier transform. 21

GAN generative adversarial network. 37, 38, 63, 120

GBCA Gadolinium-based contrast agent. 4, 5, 15, 25, 27, 28, 85, 87, 94, 95, 98, 99, 107, 108, 111, 112, 114, 118, 120–122

GM grey matter. 44, 47, 52–55, 61, 87, 89–92

GRE gradient-recalled echo. 23

ICC intra-class correlation coefficient. 50, 52, 54, 58, 61, 74, 76, 79

IDH Isocitrate dehydrogenase. 3

IQR interquartile range. 76, 89, 90, 93, 112

IR inversion recovery. 29, 31

IR-GRE inversion recovery gradient-recalled echo. 16, 23, 45

IR-SE inversion recovery-spin echo. 23, 49, 59

IRB institutional review board. 10, 132

LR logistic regression. 73

MAE mean absolute error. 45, 71

MAGiC magnetic resonance image compilation. 7, 11, 15, 85, 87–92, 94, 95, 107, 108, 110–112, 118, 122, 143

MDME multidynamic multiecho. 31

ME-SE multi echo-spin echo. 29

MPME multipathway multiecho. 31

MPRAGE magnetization prepared rapid gradient echo. 31

MR magnetic resonance. 4, 6–10, 19, 20, 23, 27, 29–31, 35, 38, 43–46, 61, 64, 69, 71, 85, 107, 111, 117, 132

MRI magnetic resonance imaging. 4–10, 15, 18, 25, 30, 31, 37, 38, 43, 61, 62, 117, 120, 121

MSE mean squared error. 48, 50, 57, 61, 73, 74, 100

NMR nuclear magnetic resonance. 18

NN neural network. 32–34, 37, 38, 120

NSE nomalized squared error. 46, 49–51

NSF nephrogenic sistemic fibrosis. 5

nWM normal white matter. 87, 143

PD proton density. 6, 8–11, 15, 19, 21–23, 30, 31, 43–46, 50–53, 57–61, 63, 70, 71, 85, 87, 89, 91, 98, 107, 110, 111, 117, 121, 122, 144

PDw PD-weighted. 10, 31, 49, 52–59, 62

post-T1 post-contrast T1. 99, 100, 102, 103

post-T1w post-contrast T1-weighted. 4, 10, 11, 15, 25, 28, 69, 72, 73, 85, 87, 92, 93, 98–101, 104, 107–111, 113, 114, 132, 144

post-T2 post-contrast T2. 100, 103

PSNR peak signal-to-noise ratio. 48, 56, 57, 73, 74, 100

ReLU rectified liner unit. 32

RF radio-frequency. 18–23, 29

ROC receiver operating characteristic. 143, 144

ROI region of interest. 46–48, 50, 55, 56, 58, 62

SD standard deviation. 74, 89, 100

SE spin echo. 23, 45, 49

SNR signal-to-noise ratio. 48, 52, 54–56, 61, 62, 121

SPGR spoiled gradient-recalled echo. 22, 29

SSIM structural similarity index. 48, 50, 56, 57, 61, 62, 73, 74, 100

STIR short-TI inversion recovery. 59

SVM support vector machine. 73

T1 longitudinal relaxation time. 6, 8–11, 15, 19–25, 27–31, 43–46, 50–53, 57–61, 63, 70, 71, 85, 87, 89, 91, 94, 95, 98–100, 102, 107, 110, 111, 117, 121, 122, 143

T1e T1-weighted contrast-enhancement. 87, 89, 90, 92–94, 108, 110, 143

T1w T1-weighted. 4, 10, 11, 15, 16, 22, 25, 27, 31, 32, 43–46, 49, 51–59, 61, 62, 69–73, 85, 87, 88, 91–95, 98–101, 104, 107, 108, 110, 112, 114, 118–122, 132, 143, 144

T1w-FLAIR T1-weighted fluid attenuated inversion recovery. 59

T2 transversal relaxation time. 6, 8–11, 15, 19–25, 27–31, 43–46, 50–53, 57–61, 63, 70, 71, 85, 87, 89, 91, 94, 95, 98, 100, 103, 107, 110, 111, 117, 119, 121, 122, 143

T2*w T2*-weighted. 10, 49, 52–59, 62, 63, 119

T2h non-enhancing T2-weighted hyperintensity. 87, 90, 92, 93, 110, 143

T2w T2-weighted. 4, 10, 11, 15, 31, 32, 43, 45, 46, 49, 51–59, 61–63, 69–80, 85, 87, 88, 92, 93, 95, 107, 108, 110, 119, 122, 132, 144

T2w-FLAIR T2-weighted fluid attenuated inversion recovery. 4, 10, 11, 31, 46, 49, 52–59, 62, 63, 69–80, 85, 87, 88, 92, 93, 107, 108, 110, 119, 132, 144

TE echo time. 7, 10, 21, 22, 29, 30, 45, 49, 59, 130, 132

TI inversion time. 7, 21, 23, 29, 30, 45, 49, 59, 132

TR repetition time. 7, 21, 22, 29, 30, 45, 49, 59, 132

VFA variable flip angle. 10, 29

ViT vision transformer. 38, 120

WHO World Health Organization. 3

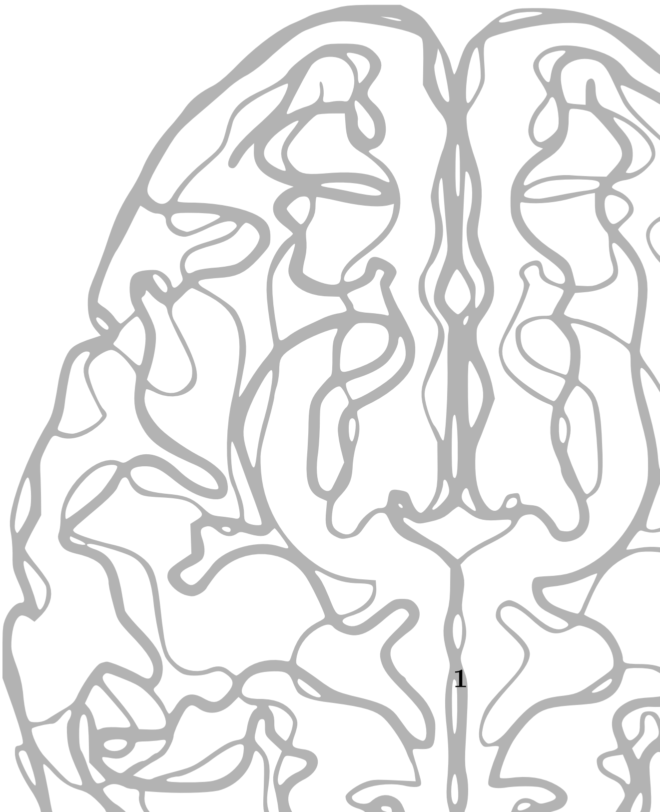
WM white matter. 43, 47, 50, 52–55, 61, 70, 87, 89–92

XGB extreme gradient boosting. 72

Contents

Part I

Introduction



1

Introduction

Contents

1.1 Motivation

Malignant glioma is the most common primary brain tumor in adults with an annual incidence of approximately 5 cases per 100.000 population (Wen and Kesari, 2008). Practically all adult-type gliomas progress with time and present an overall poor prognosis (Thust et al., 2018). This brain tumor is defined by its origin in the glial cells (Ostrom et al., 2020). According to the fifth edition of the World Health Organization (WHO) classification of tumors of the central nervous system (CNS), adult-type diffuse gliomas are an heterogeneous family, which is composed of three different types distinguished by their genetic characteristics (such as Isocitrate dehydrogenase (IDH) genes mutated versus non-mutated or wildtype) and prognostic outcomes (McNamara et al., 2022):

- **Astrocytoma, IDH-mutant:** grade 2, 3, or 4
- **Oligodendroglioma, IDH-mutant and 1p/19q-codeleted:** grade 2 or 3
- **Glioblastoma, IDH-wildtype:** grade 4

The specific tumor grade determines the treatment plan for each particular case (Horbinski et al., 2022). In general, the standard therapy includes maximal surgical resection whenever feasible, radiotherapy and chemotherapy (Thust et al., 2018; Wen and Kesari, 2008). Although maximal surgical resection tends to increase life expectancy of the patients, the infiltrative pattern of gliomas hinders the whole elimination (Wen and Kesari, 2008). Recently, several advances have been achieved in precision oncology and immunotherapy (Tan et al., 2020), but gliomas still maintain an unfavorable prognosis (Holland, 2001). The expected survival remains especially poor in glioblastoma, with approximately 40% survival in the first year after diagnosis and 17% in the second year (Thakkar et al., 2021). Thus, in those cases survival prediction is a key task for efficient treatment and surgery planning.

As for the diagnosis, the medical imaging modalities usually employed in clinical practice are magnetic resonance imaging (MRI) or computerized tomography (CT) (Wen and Kesari, 2008). Unlike CT, MRI physical principles makes it a non-invasive, innocuous and harmless imaging modality; consequently it is generally the modality of choice. In addition, magnetic resonance (MR) images present an excellent contrast between soft tissues that facilitates the distinction of brain tumors. The images that are typically acquired during an MR scan are of a qualitative nature since they are based on relative differences between tissues. Thus, they are known as **qualitative or weighted MRI** (Gulani and Seiberlich, 2020).

MR weighted images are acquired by applying specific pulse sequences with a combination of acquisition parameters, which are selected by the operator. The choice of both sequence and parameters gives rise to a specific weighting of the tissue magnetic properties. All the tissue magnetic properties always affect the weighted images, although a predominant weighting of a specific magnetic property generally exists. Each image weighting provides complementary information for diagnosis (Lu et al., 2005), since different structures and/or conditions are more clearly visible in each of them. Thus, an MRI scan protocol typically consists of a number of sequences for obtaining various weighted images. These weighted images are routinely used by radiologists for diagnosis through visual inspection.

Specifically, the standard protocol for glioma assessment, portrayed in Figure 1.1, generally includes (at least) a T1-weighted (T1w), T2-weighted (T2w), T2-weighted fluid attenuated inversion recovery (T2w-FLAIR), and post-contrast T1-weighted (post-T1w) (Ellingson et al., 2015a; Thust et al., 2018). The latter deserves an special mention; it is obtained after the injection of a Gadolinium-based contrast agent (GBCA) and it is key due to its ability to reveal the impairment in the blood brain barrier (BBB). Following the visual inspection criteria for diagnosis, the T1w and post-T1w are visually compared seeking the potential contrast enhancement in the latter caused by the extravasation of GBCA into the perivascular space (Hattingen et al., 2017). Evaluating the enhancement is crucial for tumor resection planing; literature suggests that in patients without comorbidities a resection above 90% of the tumor contrast enhancement improves patient outcome and minimize recurrence (Brown et al., 2016; Gilard et al., 2021; Lacroix et al., 2001). Also, the presence of enhancement is associated with aggressive tumor behavior in high-grade

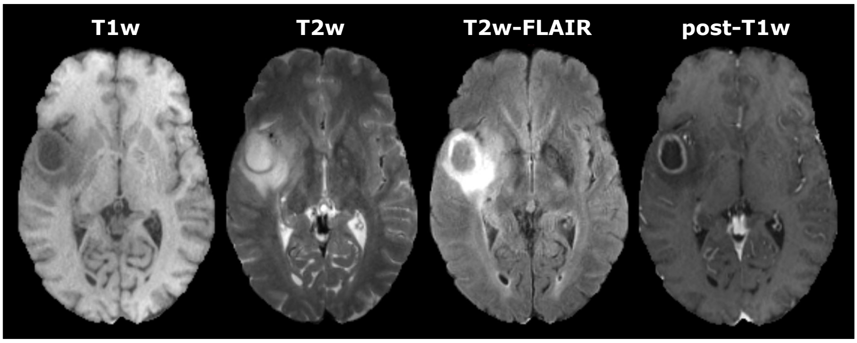


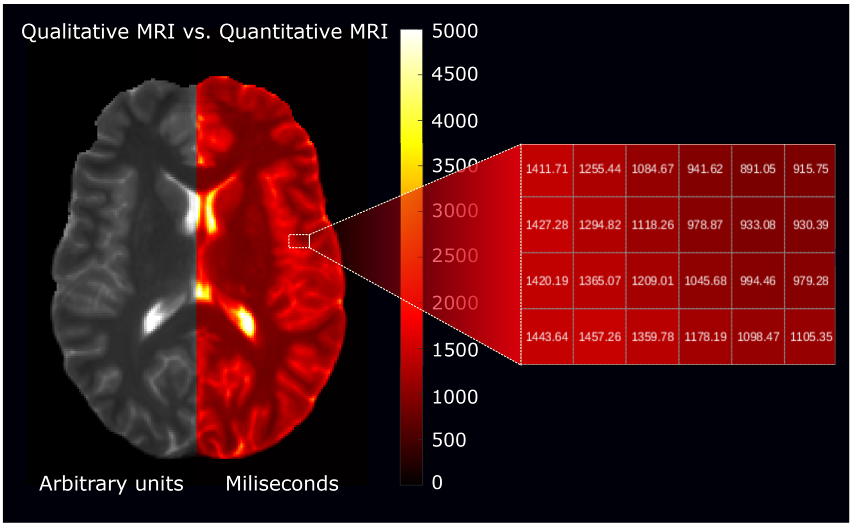
Figure 1.1: Weighted images included in a standard MRI protocol for glioma assessment. These image modalities are usually acquired, but other modalities could be also included depending on the institution. See how each image modality is able to highlight different properties of the tumor.

gliomas and determines the treatment plan (Warntjes et al., 2018).

However, this diagnostic procedure presents two major issues:

a) The arbitrary scale of weighted images hinders the usage of quantitative diagnostic methods. Nevertheless, these quantitative methods could successfully complement the standard diagnosis procedure for distinguishing different post-treatment conditions and understanding biological changes within the tumor, among others (Lescher et al., 2015; Pirkl et al., 2021). Hence, some procedures are difficult to be performed only through visual inspection of the weighted images, such as the detection of tumor infiltration into the peritumoral edema (Blystad et al., 2017).

b) GBCAs are also problematic by their own nature; although GBCAs are generally deemed safe, between 0.07 to 2.4% of injected patients suffer from mild adverse reactions and a lower rate (around 0.03%) from severe complications (Forghani, 2016; Granata et al., 2016; Ramalho et al., 2016). Severe life-threatening anaphylactic type reactions are exceedingly rare (0.001% to 0.01%) (Forghani, 2016; Granata et al., 2016). In addition, nephrogenic sistemic fibrosis (NSF) is a rare but serious side effect, which can arise in people with severe kidney problems (Grobner, 2006). Other concern among the community is the possible deposition of GBCAs in several tissues in the brain (Gulani et al., 2017; Kanda et al., 2014; Runge, 2016), especially in patients who have to undergo several follow-up acquisitions with GBCAs such as the oncological patient. Also, the environmental issues associated with the presence of GBCAs in wastewater (Inoue et al., 2020; Rogowska et al., 2018). Finally, the usage of GBCAs results in patient discomfort during intravenous injection and increases scan time, the need of skilled manpower, and costs (Shankar et al., 2018).



1411.71	1255.44	1084.67	941.62	891.05	915.75
1427.28	1294.82	1118.26	978.87	933.08	930.39
1420.19	1365.07	1209.01	1045.68	994.46	979.28
1443.64	1457.26	1359.78	1178.19	1098.47	1105.35

Figure 1.2: Differences between qualitative MRI (i.e., weighted images) and quantitative MRI (i.e., parametric maps). The intensity of weighted images does not present units, whereas parametric maps present an absolute scale, measured in milliseconds in the T1 map shown in the image.

An alternative option to perform the diagnosis is to find the tissue magnetic properties themselves, methodology known as **quantitative MRI**. These properties include the longitudinal relaxation time (T1), transversal relaxation time (T2), and proton density (PD) — jointly referred to as parametric maps —. Parametric maps present an intrinsic quantitative nature and absolute scale (see Figure 1.2). They are also more robust than weighted images due to their lower sensitivity to MRI hardware (Weiskopf et al., 2013). In addition, these parameters play an important role in tissue characterization in healthy and diseased stages of pathologies such as epilepsy (Conlon et al., 1988) or multiple sclerosis (Larsson et al., 1989) and, also, for tumor detection (Yankeelov et al., 2011). Regardless of the aforementioned advantages of parametric maps compared to weighted images, radiologists are not used to perform diagnosis solely on the basis of these maps (Gulani and Seiberlich, 2020; Hagiwara et al., 2017).

Synthetic MRI serves as a bridge between qualitative and quantitative MRI (Ji et al., 2020). Although this concept has recently gain popularity, it was first defined by Bobman et al. (1985) as a three-step procedure:

1. Acquisition of a set of MR weighted images with different operator-selectable acquisitions parameters.
2. Computation of the quantitative parametric maps (i.e., T1, T2, PD) from the acquired weighted images by voxel-wise fitting a known relaxation model.
3. Synthesis of weighted images using the theoretical equations that describe the per-pixel MR image intensity as a function of operator-selectable acquisition

parameters, such as echo time (TE), repetition time (TR), inversion time (TI), etc., and the aforementioned parametric maps.

Thus, this low-cost procedure could retrospectively enhance patient throughput and facilitate radiologists routine. On the one side, parametric maps could enable the quantification of subtle changes within the tissues. On the other side, artifacted weighted images could be replaced by their synthesized counterversion without the need of re-acquistion and/or some image modalities could be directly not acquired, which imply a reduction of the protocol time. Recently, other approaches have been proposed that learn the mapping between different pairs of MR weighted images (Chartsias et al., 2017; Dar et al., 2019; Sohail et al., 2019). Although these methods have been sometimes referred to as Synthetic MRI in the literature, we believe that medical image translation is a more descriptive name for them. Thus, throughout this Thesis dissertation we only consider Synthetic MRI methods those that fit within the three-step definition given above. This consideration is done since these methods are flexible in regards to the MR images that can be synthesized.

Step 2 (i.e., the estimation of parametric maps from a set of acquired weighted images) has been traditionally carried out with techniques known as relaxometry. However, their extremely long acquisitions, which are unfeasible in clinical practice, hamper their inclusion in routine protocols. Nowadays, fast multiparametric mapping techniques, such as MR Fingerprinting (Ma et al., 2013) or magnetic resonance image compilation (MAGiC) (Warntjes et al., 2008), have taken the stage. However, these techniques, albeit faster than traditional relaxometry, still require a not negligible acquisition time difficult to accommodate in an already tight protocol. Moreover, neither their sophisticated sequences nor their associated commercial software are commonly available worldwide, thereby their practical use is not widespread (Ji et al., 2020).

Parametric mapping could also be boosted by **deep learning (DL)** techniques if the computation could be performed from only conventional weighted images, which are part of any routine protocol. The avoidance of *ad-hoc* pulse sequences may facilitate the availability and consolidation of these approaches in clinical care. However, training DL methods is usually a data-demanding task (Isola et al., 2017) and, to the best of our knowledge, there are no public databases that include several types of weighted images together with their corresponding parametric maps. In addition, obtaining these parametric maps with in-house acquisitions is also difficult due to sensible time restrictions that ethic committees (and common-sense) impose. In this context, finding solutions that facilitate the training of DL models with few representative cases is mandatory. Furthermore, in order for the proposed approaches to be useful, the diagnostic value of the DL-synthesized parametric maps and weighted images should be validated with clinical endpoints (Gulani and Seiberlich, 2020).

In this Thesis we propose Synthetic MRI approaches for the computation of parametric maps and the synthesis of several types of weighted images, which operate using as inputs conventional weighted images. The Synthetic MRI approaches are based on DL, and some solutions for training the models with small databases are also presented. Finally, the utility of the proposed approaches is validated in different clinical applications for brain tumor diagnosis. Thus, we show that the DL synthesized parametric maps and weighted images can be used as a surrogate of those acquired with traditional techniques or through commercial software.

1.2 Objectives

The main objective of this Thesis is **to enhance the diagnosis of brain tumors using a Synthetic MRI paradigm, i.e., by means of the computation of T1, T2, and PD parametric maps, as well as the synthesis of several weighted images, out of routine sequences typically used in clinical protocols.** Hence, the methods should be effortless, i.e., should not add up any extra time to the acquisition protocols and should not require any non-conventional MR acquisition sequence. To achieve this objective, DL techniques will be employed.

The main objective can be itemized into the following sub-objectives:

- **O1:** To propose and develop a Synthetic MRI approach for T1, T2, and PD multiparametric mapping from routine MRI acquisitions commonly used in clinical practice and widely available in all the scanners.
- **O2:** To synthesize different modalities of MR weighted images from the computed parametric maps. These images not only have to be visually similar to those acquired, but they also have to show similar quantitative characteristics so that they can be used by quantitative and automatic diagnostic methods.
- **O3:** To determine the utility of the synthesized weighted images to feed a radiomics system for predicting the expected survival of glioblastoma patients.
- **O4:** To determine the feasibility of the parametric maps computed with DL as surrogate of those obtained with relaxometry or other commercial techniques in a clinical application for predicting brain tumor T1w-enhancement without contrast agents.
- **O5:** To extend the proposed Synthetic MRI approaches for the computation of post-contrast parametric maps and the synthesis of post-contrast weighted images for automatic quantification of tumor enhancement.

1.3 Methodology

The methodology employed in this Thesis follows the engineering research approach outlined by Adrion (1993) and refined by Glass (1995). This research method involves observing existing solutions, proposing improved approaches, developing these approaches, and then iteratively measuring, analyzing, and repeating the process until further improvements are unattainable. The specific methodological phases followed in this Thesis are detailed below:

1. The informational phase: this phase entails reviewing the state-of-the-art solutions related to a specific problem, accomplished through an exhaustive literature survey. In the scope of this Thesis, we detected that the qualitative nature of weighted images could hinder the usage of quantitative diagnostic methods, particularly in neuro-oncology for brain tumor diagnosis. These methods could indeed be relevant to better characterize subtle changes within the tumors and, in turn, to improve patient outcome. However, radiologists are not used to relying only on quantitative image modalities. In this context, Synthetic MRI stands out as a solution since it serves as a bridge between qualitative and quantitative MRI. However, it is usually performed with specific sequences and commercial softwares, which limits its widespread usage.

2. The propositional phase: this phase comprises the formulation of a hypothesis, method, algorithm, or theory. In the context of the current Thesis, we proposed to develop a Synthetic MRI approach from only conventional acquisitions. To this end, DL might be a potential solution. Our final practical goal is to broaden the availability of Synthetic MRI to improve brain tumor diagnosis.

3. The analytical phase: this phase involves analyzing and validating the proposition. Regarding the computation of T1, T2, and PD parametric maps from a small number of routine sequences, we examined the suitability of convolutional neural network (CNN) architectures for addressing this type of problem. Limited large and public datasets of parametric maps was a drawback, which hinders the usage of DL to this end. Thus, we considered the possibility of performing part of the training with synthetic data.

4. The evaluative phase: this stage entails the evaluation of the proposition by means of experimentation. We validated the proposed approaches with both synthetic data and real MR acquisitions of healthy volunteers and patients diagnosed with brain tumors. In addition, public datasets, in-house acquisitions, and data obtained from collaborations with other institutions and hospitals are employed. Thus, the generalization capability of the proposed methods were also tested with different multi-center, multi-vendor acquisitions. The proposed approaches were validated by means of visual assessment, quantitative quality metrics, as well as quantifying the performance of the computed parametric maps and synthesized weighted images in different clinical applications for tumor diagnosis.

Additionally, in the course of this Thesis, we executed an additional phase, which is referred to as the **Dissemination Phase**. In this phase, we shared our motivation, final propositions, methodologies, and outcomes with the community. The dissemination was performed not only through specialized international journals and conferences, but also through newspapers and radio broadcast in order to reach both international research and local communities.

1.4 Materials: Datasets

The MR acquisitions used in this Thesis were collected with the approval of the corresponding institutional review board (IRB) and after the volunteers or patients had signed an informed written consent. The different datasets used throughout this Thesis are described below:

- ***Multicontrast Brain-MRI***: eight subjects, suspected of early Alzheimer disease, (mean age, 71.38 years ± 6.91 [SD]; 5 female, 3 male) were scanned at the University of Valladolid, Valladolid, Spain on a 3T Achieva (Philips, Best, The Netherlands). Each study was composed of four acquisitions for obtaining a total of five structural weighted images. These weighted images are T1w, PD-weighted (PDw), T2w, T2w-FLAIR, and T2*-weighted (T2*w). Details about the acquisition parameters can be found in Table A.1.

- ***Relaxometry Brain-MRI***: five healthy volunteers (mean age, 34.40 years ± 12.18 [SD]; 2 female, 3 male) were included in this dataset. The dataset was collected at the University of Valladolid, Valladolid, Spain in 2021 with a 3T Achieva (Philips, Best, The Netherlands). The protocol included relaxometry sequences for the estimation of T1, T2, and PD maps, and also T1w and T2w images. A T2w-FLAIR image with the same acquisition parameters than in previous dataset *Multicontrast Brain-MRI* was also acquired but only used for registration purposes. The T1 maps were estimated from a variable flip angle (VFA) sequence with 11 different flip angles using NOVIFAST (Ramos-Llordén et al., 2018), and the T2 maps were obtained from a multi-echo sequence with six different TEs using a least squares fit. Details about the acquisition parameters can be found in Table A.2.

- ***Multicontrast Glioblastoma***: 24 patients diagnosed with glioblastoma (mean age, 57.0 years ± 13.0 [SD]; 10 female, 14 male) were acquired at the Hospital Universitario 12 de Octubre, Madrid, Spain on a 1.5T scanner (General Electric, Waukesha, WI, USA). For each patient, four MR structural weighted images were collected, namely T1w, T2w, T2w-FLAIR, and post-T1w. Details about the acquisition parameters can be found in Table A.3.

- ***UPenn-GBM***: 611 patients diagnosed with *de novo* glioblastoma are included in this public dataset of the University of Pennsylvania, Philadelphia, Pennsylvania. The acquisitions were performed from 2006 to 2018. For 41 of these patients a follow-up MRI acquisition prior to a second resection is also available. Details about the dataset can be found in Bakas et al. (2022). This dataset includes four structural weighted images (T1w, T2w, T2w-FLAIR, and post-T1w). The images were preprocessed by resampling them to an isotropic resolution of 1mm^3, and skull-stripping. The segmentation into three different regions — necrosis, enhancement, and edema — were also available. Different subgroups of patients from this dataset were employed in this Thesis. Specifically, in ***Upenn-GBM-A*** (used in Chapter 5) we selected 493 patients (mean age, 61.94 years ± 11.98 [SD]; 91 female, 129 male) and in ***Upenn-GBM-B*** (used in Chapter 6) we selected 220 patients (mean age, 63.15 years ± 12.23 [SD]; 197 female, 296 male). In both cases the acquisitions were performed with a 3T Magnetom Trio (Siemens Healthcare, Erlangen, Germany) and with homogeneous acquisition parameters. Details about the acquisition parameters can be found in Table A.4.

- ***GLIOMA***: 15 patients (mean age, 39.33 years ± 10.40 [SD]; 6 female, 9 male) with different grades of gliomas, who were scanned at the Erasmus MC, Rotterdam, The Netherlands. The patients had undergone tumor resection before the acquisition and were scanned between 2018 and 2020 with a 3T Sigma Premier (General Electric, Waukesha, WI, USA). The image modalities included are T1w, T2w, T2w-FLAIR, post-T1w, and, also, pre-contrast MAGiC for T1, T2, and PD parametric mapping. In one patient the T2w-FLAIR was not acquired due to a protocol deviation. Details about the acquisition parameters can be found in Table A.5.

Regarding the hardware resources, they vary depending on the experiment. The main workstation includes a CPU Intel(R) Core(TM) i7-4790 CPU @ 3.60GHz and 16 GM RAM memory. In addition, throughout the development of this Thesis, different GPUs were employed, which are listed below:

- NVIDIA GeForce RTX 2080 Ti
- NVIDIA Quadro RTX 6000
- NVIDIA GeForce GTX 1070

1.5 Publications

Below, the list of publications related to this Thesis is included. These publications are also schematically shown in Figure 1.3. Each contribution is related with, at least, one of the aforementioned objectives although more than one publication can be related with the same objective.

- Core publications of this Thesis:
  - Publications in indexed international journals:
    - **Moya-Sáez, E.**, Peña-Nogales, Ó., de Luis-Garcia, R., and Alberola-López, C. A deep learning approach for Synthetic MRI based on two routine sequences and training with synthetic data. *Computer Methods and Programs in Biomedicine.* 2021; 210, 106371.

    - **Moya-Sáez, E.**, Navarro-González, R., Cepeda, S., Pérez-Núñez, A., de Luis-Garcia, R., Aja-Fernández, S., and Alberola-López, C. Synthetic MRI improves radiomics-based glioblastoma survival prediction. *NMR in Biomedicine.* 2022; 35(9), e4754.

    - **Moya-Sáez, E.**, de Luis-Garcia, R., Nunez-Gónzalez, L., Alberola-López, C., and Hernández-Tamames, J.A. Brain tumor T1w-enhancement prediction from pre-contrast conventional weighted images using synthetic multiparametric mapping and deep learning. *Submitted.*

- Publications in non-indexed international journals:
    - **Moya-Sáez, E.**, de Luis-Garcia, R., and Alberola-López, C. Toward deep learning replacement of gadolinium in neuro-oncology: A review of contrast-enhanced Synthetic MRI. *Frontiers in Neuroimaging.* 2023; 2, 1055463.

- Publications in international conferences:
    - **Moya-Sáez, E.**, Peña-Nogales, Ó., Sanz-Estébanez, S., de Luis-Garcia, R., and Alberola-López, C. CNN-based synthesis of T1, T2 and PD parametric maps of the brain with a minimal input feeding. *ISMRM & SMRT Virtual Conference & Exhibition*, Virtual, August 2020; 3806.
    - **Moya-Sáez, E.**, de Luis-Garcia, R., and Alberola-López, C. A self-supervised deep learning approach to synthesize weighted images and T1, T2, and PD parametric maps based on MR physics priors. *ISMRM & SMRT Annual Meeting & Exhibition, An Online Experience*, Virtual, May 2021; 2169.
    - Navarro-González, R., **Moya-Sáez, E.**, de Luis-Garcia, R., Aja-Fernández, S., and Alberola-López, C. Synthetic MRI aids in glioblastoma survival prediction. *Joint ISMRM-ESMRMB & SMRT 31st Annual Meeting*, London, England, UK, May 2022; 3928.
    - **Moya-Sáez, E.**, de Luis-Garcia, R., Hernández-Tamames, J.A., and Alberola-López, C. Pre and Post contrast Simultaneous Parametric Mapping of Glioblastomas from routine T1 weighted images for Quantitative Enhancement Assessment. *2023 ISMRM & SMRT Annual Meeting & Exhibition*, Toronto, Canada, June 2023; 2352.
    - **Moya-Sáez, E.**, de Luis-Garcia, R., Alberola-López, C., and Hernández-Tamames, J.A. Deep learning-based post-contrast imaging free of exogenous contrast agents. *3rd Annual Meeting of the ISMRM Iberian Chapter*, Valladolid, Spain, July 2023; P1.13. Best poster paper award.
    - **Moya-Sáez, E.**, Nunez-Gónzalez, L., de Luis-Garcia, R., Alberola-López, C., and Hernández-Tamames, J.A. Post-contrast multi-parametric mapping from only pre-contrast conventional weighted images. *2024 ISMRM & ISMRT Annual Meeting & Exhibition*, Singapore, May 2024; 2809.

- Invited talks:
    - **Moya-Sáez, E.**, de Luis-Garcia, R., and Alberola-López, C. Synthetic image in MRI *36 Conference Sociedad Española de Radiología*

Médica (SERAM) / XXXI Conference Colegio Interamericano de Radiología (CIR), Málaga, Spain, May 2022.

In the subsequent list we also include other contributions accomplished during the progress of this Thesis but not directly related with the main scope of it.

- Other publications accomplished during the progress of this Thesis:
    - ○ Publications in indexed international journals:
        - – Martín-González, E., **Moya-Sáez, E.**, Menchón-Lara, R.M., Royuela-del-Val, J., Palencia-de-Lara, C., Rodríguez-Cayetano, M., Simmross-Wattenberg, F., and Alberola-López, C. A clinically viable vendor-independent and device-agnostic solution for accelerated cardiac MRI reconstruction. *Computer Methods and Programs in Biomedicine*. 2021; 207, 106143.
    - ○ Publications in international conferences:
        - – Rodríguez-Galván, J.R., Martín-Martín, C., **Moya-Sáez, E.**, Tristán-Vega, A., Aja-Fernández, S., and Alberola-López, C. DL Diffusion MRI enhancing may lead to incorrect diagnosis. *2nd Annual Meeting of the ISMRM Iberian Chapter*, Lisbon, Portugal, June 2022; P2.15.

1.6 Document overview

This document is divided in five parts; Part I includes the introduction and the context of this Thesis dissertation. In Part II the baseline Synthetic MRI approach from routine sequences is presented. This baseline method is extended in Parts III and IV to be used in specific clinical applications for brain tumor diagnosis. Specifically, the Synthetic MRI approach is extended in Part III for survival prediction of glioblastoma patients and in Part IV for brain tumor T1w-enhancement prediction. Finally, Part V includes some final remarks and future work.

A more detailed description is outlined next:

- **Part I:** Introduction part, which includes the introduction, motivation and context of this Thesis dissertation.
    - ○ **Chapter 1** introduces and motivates this Thesis dissertation. The main goal pursued in the Thesis as well as the specific sub-objectives are enumerated. Also, the methodology and materials utilized are explained. Finally, we include a list of journal and conference publications achieved during the realization of this Thesis.
    - ○ **Chapter 2** includes background material required for a complete understanding of the topics addressed in this Thesis. An overview of the state-of-the-art related with this Thesis is also incorporated.

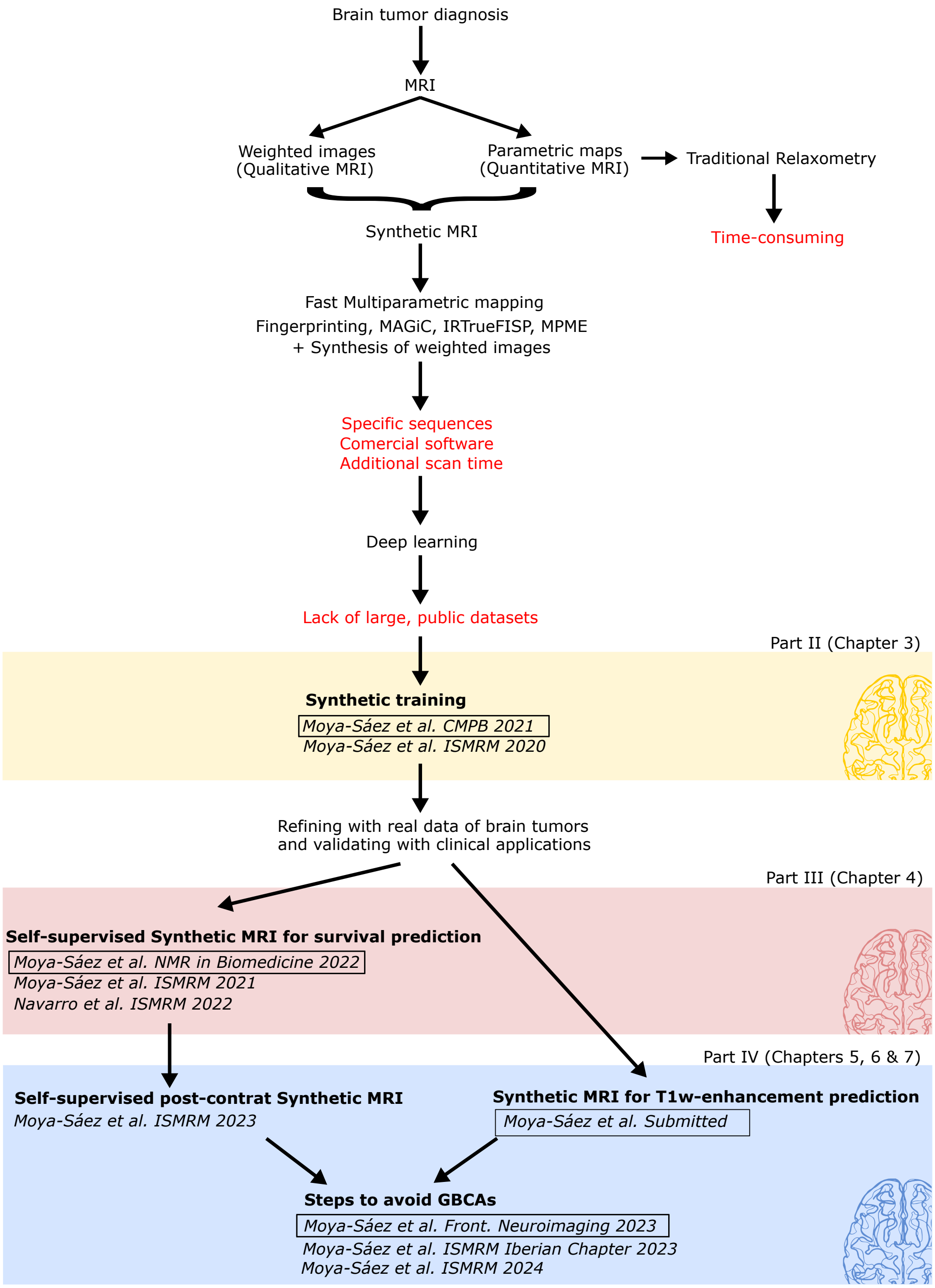


Figure 1.3: Outline of the motivation and contributions of this Thesis. Red text refers to the drawbacks found in the state-of-the-art and limitations that are overcome through this work. Publications related with this Thesis are also included; those displayed inside a square correspond to the ones published in international journals.

- **Part II:** This part proposes the Synthetic MRI approach for obtaining parametric maps from conventional weighted images acquired with routine sequences. This part consists of **Chapter 3**, in which we present the Synthetic MRI approach for the DL computation of T1, T2, and PD parametric maps from only a pair of weighted images (a T1w and a T2w) acquired with routine sequences. The training of the DL method is performed mainly with synthetic data in order to overcome the lack of large, public datasets with parametric maps. The synthesis of several types of weighted images from the computed parametric maps is also performed.

- **Part III:** This part proposes an extension of the Synthetic MRI approach to improve a radiomics system for survival prediction in glioblastoma. This part consists of **Chapter 4**, in which we introduce a self-supervised extension of the Synthetic MRI approach described in Chapter 3. This extension allows us to perform the training of the DL network with acquired weighted images of glioblastoma patients. Thus, reference parametric maps are not needed for training. The synthesized weighted images are next employed to feed a radiomics system for predicting the expected survival of such patients. Both the radiomics system and image synthesis method are described in this chapter. The results show how the synthesized images can replace one of the acquired sequences in this clinical application.

- **Part IV:** This part contributes to improve the current procedures for the detection and quantification of brain tumor T1w-enhancement.

    - **Chapter 5** describes a refinement of the Synthetic MRI approach described in Chapter 3 by training it with a small dataset of parametric maps of glioma patients obtained with MAGiC. The DL-computed parametric maps are employed for predicting T1w-enhancement without contrast agents, and those maps are proved to be useful in that application. Thus, this chapter shows how DL-synthesized parametric maps could be a surrogate of those maps obtained with relaxometry or commercial software.

    - **Chapter 6** includes an extension of the self-supervised method previously described in Chapter 4 for the computation of pre- and post-contrast parametric maps. In addition, this chapter shows how those maps make it possible to automatically quantify tumor T1w-enhancement as a complementary tool to standard visual assessment.

    - **Chapter 7** proposes a cascade of two CNNs for pre- and post-contrast parametric mapping and the synthesis of post-T1w images from only two pre-contrast routine weighted images. Thus, preliminary steps towards the avoidance of GBCAs are presented in this chapter.

- **Part V:** Conclusions part, which includes the conclusions and final remarks. This part consists of **Chapter 8**, in which the main contributions, advantages, and limitations are included. Furthermore, we figure out some future lines that could arise from this research.

- **Part VI:** Appendices part, which includes supplementary material of this Thesis dissertation.
    - **Appendix A** includes details about the acquisition parameters for each dataset employed in this Thesis.
    - **Appendix B** derives the inversion recovery gradient-recalled echo (IR-GRE) theoretical pulse sequence equation used in this Thesis.
    - **Appendix C** includes additional explanations about the radiomics system for glioblastoma survival prediction employed in Chapter 4.
    - **Appendix D** includes additional explanations about the voxel-wise statistical classifications of T1w-enhancement employed in Chapter 5.

2

Background

Contents

2.1 Principles of Magnetic Resonance Imaging

2.1.1 From spins to signal

The basic principle of MRI is rooted in the nuclear magnetic resonance (NMR) phenomenon described by Bloch (1946) and Purcell et al. (1946). Nuclei with unpaired protons or neutrons possess a property called spin, which makes them “MR active”. Several nuclei, such as ^{1}9F, ^{1}3C, ^{2}3Na, etc., present this characteristic, but the one most commonly used in MRI is ^{1}H due to its presence in water molecules. ^{1}H is a single, positively charged proton. The NMR phenomenon consists of the interaction of a nuclear spin in the presence of an external magnetic field $\mathbf{B}_0$. Thus, this phenomenon can be explained with both quantum mechanics and Newtonian theories. From a quantum mechanics explanation, protons can occupy multiple energy levels, so they align either parallel or antiparallel to the $\mathbf{B}_0$ field. At room temperature, a small excess of parallel protons exists and thus the net magnetization vector ($\mathbf{M}$) is aligned with the $\mathbf{B}_0$ direction, which conventionally has been fixed to the longitudinal ‘z’-axis. The net magnetization vector at this steady state situation is reffered to as $\mathbf{M}_0$. Hereinafter, magnetic moments will be explained in purely Newtonian terms in order to simplify the explanation.

The $\mathbf{B}_0$ field produces a secondary spin called precession that causes $\mathbf{M}_0$ to produce a circular path around $\mathbf{B}_0$. The rate at which $\mathbf{M}_0$ precesses around the external magnetic field is known as the the precessional frequency (ω_0) and is determined by the Larmor equation:

$$\omega_0 = \gamma ||\mathbf{B}_0||_2 [\text{rad/s}], \tag{2.1}$$

where γ is the gyromagnetic ratio, which is the precessional frequency for an specific nucleus at 1T. Specifically, for ^{1}H nuclei it is 42.57 MHz/T.

The magnetic moments of the spinning nuclei are perturbed by means of time-varying radio-frequency (RF) pulses. The RF fields ($\mathbf{B}_1(t)$) are applied in the transverse plane (i.e., ‘x’-‘y’ plane) and rotate at the Larmor frequency. These fields cause the ^{1}H nuclei to resonate, that is to absorb energy from the RF pulse, which produces another rotation to the net magnetization vector ($\mathbf{M}$) modelled by:

$$\frac{\partial \mathbf{M}}{\partial t} = \mathbf{M} \times \gamma(\mathbf{B}_0 + \mathbf{B}_1(t)) \tag{2.2}$$

The new precession around the $\mathbf{B}_1(t)$ magnetic field tips the magnetization vector away from the longitudinal direction in a spiral way towards the transverse plane. The amount of rotation from the longitudinal axis ‘z’ is fixed by the intensity and duration of the $\mathbf{B}_1(t)$, and is accounted for as the flip angle (α). Thus, the application of any RF pulse is considered to be an excitation of the system.

For signal measurement, a receiver coil is situated in the transverse plane. As the net magnetization vector rotates at the Larmor frequency around the transverse plane, it induces a voltage in the receiver coil according to Faraday's law of electromagnetic induction (i.e., a voltage will be induced in any conductor exposed to a changing magnetic field). It is important to note that only the transverse component of **M** induces a voltage. This voltage corresponds to the MR signal, whose magnitude is proportional to the PD.

After a certain time, the RF pulse is switched off. This causes that an increasingly higher amount of nuclei become out of phase with each other and the signal induced in the receiver coils begins to decrease. This process is call free induction decay (FID). As a consequence, the net magnetization returns to its resting state by restoring it back into the longitudinal axis in a phenomenon called relaxation. The relaxation rate is different in each tissue, and this is why relaxation forms the basis of tissue contrast. Two process are involved:

- **Longitudinal relaxation**: this is the process by which the net magnetization tends to recover its original value on the longitudinal axis (see Figure 2.1). It is caused by the interaction between the spins and the surrounding medium, i.e., spin-tissue interactions. This process is driven by parameter T1 and it is also known as T1 recovery.
- **Transversal relaxation**: this is the process by which the net magnetization tends to fade away (see Figure 2.1). It is caused by the interactions between the micro magnetic fields generated by neighboring spins, i.e., spin-spin interactions, and the interaction between spins and the field inhomogeneities of the external magnetic field. This process is driven by parameter T2 and it is also known as T2 decay. In the presence of field inhomogeneities the T2 parameter decreases to the so-called T2*.

If we include the relaxation process in the previously stated Eq. 2.2, we obtain the well-known Bloch equation that models the whole MR dynamics:

$$\frac{\partial \mathbf{M}}{\partial t} = \mathbf{M} \times \gamma \mathbf{B} - \frac{M_x \mathbf{i} + M_y \mathbf{j}}{T_2} - \frac{(M_z + M_0)\mathbf{k}}{T_1}, \tag{2.3}$$

where M_z is the longitudinal magnetization, M_0 is the equilibrium magnetization, and M_x and M_y are both components of the transverse magnetization. At this stage, a common approach involves eliminating the effect of the Larmor precession by defining a rotating frame of reference $[x', y', z']$ synchronized with the precession frequency of the spins. Thus, the equations describing both relaxation phenomena can be independently described in the rotating frame of reference by:

$$\begin{cases} \frac{\partial M_{z'}}{\partial t} = -\frac{M_{z'} + M_0}{T_1} \\ \frac{\partial M_{x'y'}}{\partial t} = -\frac{M_{x'y'}}{T_2}, \end{cases} \tag{2.4}$$

where the longitudinal relaxation is described by:

$$M_{z'}(t) = M_0 + (M_{z'}(0) - M_0)e^{-\frac{t}{T_1}}, \tag{2.5}$$

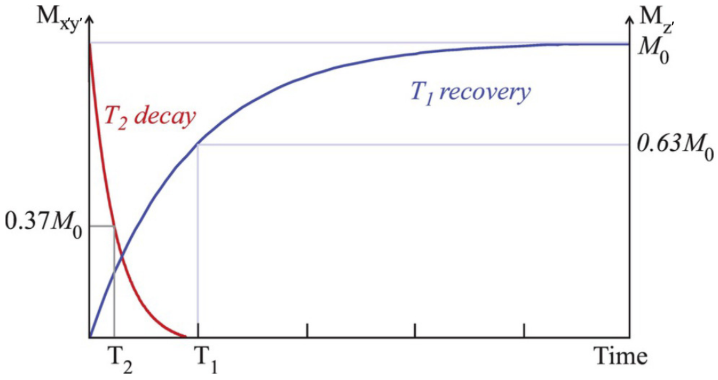


Figure 2.1: Longitudinal relaxation (or T1 recovery) and transversal relaxation (or T2 decay) after a 90° RF pulse. (Figure taken from McRobbie et al. (2017)).

and the transverse relaxation is described by:

$$M_{x'y'}(t) = M_{x'y'}(0)e^{-\frac{t}{T2}} \tag{2.6}$$

2.1.2 From signals to images

The formation of MR images requires a spatial information encoding of the induced MR signals (Lauterbur, 1973). Spatial magnetic gradient fields are used to this end. First, it is necessary to selectively excite an specific imaging plane (called slice) whose thickness is in the order of millimeters. To achieve this, a spatial magnetic field gradient is employed, which generates a magnetic field varying with the position. Consequently, the precessional frequency of the nuclear spins becomes position-dependent. Thus, a spatial field gradient perpendicular to the slice of interest has to be applied and, at the same time, the frequency of the B1 pulse has to be adjusted to match the precessional frequency of that specific slice. The most common plane orientations are axial (plane 'x'-'y'), coronal (plane 'x'-'z'), and sagittal (plane 'y'-'z'), although other oblique orientations are also possible. Hence, a stack of 2D slices are sequentially acquired in order to scan the whole object, obtaining a so-called multi-slice acquisition.

Second, it is necessary to code signals induced for each spatial location of the excited slice by means of imaging encoding gradients. Were this not done, the MR signal induced in the receiver coil would correspond to the aggregation of the signal generated by all the spins. Thus, it would not be possible to distinguish between the signals induced for each spatial location. Hence, to create an image of, for example, an axial slice, encoding gradients have to be applied in both the 'x' and 'y' directions. These gradients are known as frequency and phase encoding gradients. By ingeniously combining these encoding gradients, we can take advantage of the precessional frequency changes they induce along the direction of the encoding gradient and the phase accrual they generate. Ultimately, this process codes the signal within the entire slice in both frequency and phase increments, facilitating subsequent image reconstruction.

The signal measured by the receiver coil ($s(t)$) is the ensemble of all signals from all the spins excited within the slice.

$$s(t) = \int_x \int_y m(x,y) e^{-i2\pi[k_x(t)x + k_y(t)y]} \partial x \partial y, \tag{2.7}$$

where $m(x,y)$ is the image to be reconstructed, and

$$k_x(t) = \frac{\gamma}{2\pi} \int_0^t G_x(\tau) \partial\tau, \tag{2.8}$$

$$k_y(t) = \frac{\gamma}{2\pi} \int_0^t G_y(\tau) \partial\tau, \tag{2.9}$$

are the k-space trajectories along each axis. Eq. 2.7 corresponds to the 2D Fourier transform (FT), which connects the image space with the signal measured in the k-space. Therefore, the connection between the image and the sampled data (k-data) is determined by the trajectory defined by the imaging encoding gradients. Alternatively, the spatial location or encoding concept could be extended from a 2D slice to a 3D volume. Consequently, the object would be encoded across three dimensions using the appropriate k-space trajectory $\mathbf{k}(t) = [k_x(t), k_y(t), k_z(t)]$, resulting in a 3D FT. It is worth mentioning that both the k-space and the image space signals are complex, albeit the final form of the image space usually is the magnitude image after discarding its phase information. Subsequently, if the k-space is fully acquired with a Cartesian trajectory, the reconstruction of the final image can be performed with a discrete Fourier transform (DFT).

2.1.3 Fundamental MR pulse sequences

MR image acquisitions are performed by means of pulse sequences. A pulse sequence consists in a specific combination of RF pulses, gradients applications, and, also, the periods of time in between. Thus, the purposes of pulse sequences are: (a) to modify the phase of spins and therefore produce a signal or echo, which can be measured in the receiver coil, and (b) to enable the manipulation of some acquisition parameters to produce different types of image contrasts.

An image has a contrast when it includes regions with both high and low signal intensities, along with some areas displaying intermediate signal levels. Tissues with large transverse magnetization at measurement time present high signal (appearing white or hyperintense), while those with small transverse magnetization have low signal (appearing black or hypointense). The contrast in an image is influenced by intrinsic contrast mechanisms (T1, T2, and PD, among others) and extrinsic contrast acquisition parameters, such as TR, TE, TI, and flip angle. The former are inherent to the tissue being imaged, whereas the latter are under the control of the system's operator during the configuration of the pulse sequence. Thus, these acquisition parameters are selected in order to weight the image towards one intrinsic contrast mechanism and away from the others, thereby, they determine the weighting of the acquired image.

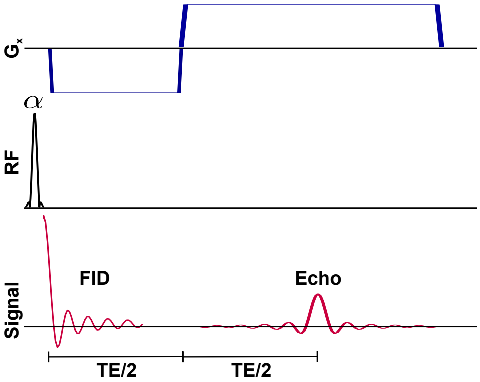


Figure 2.2: Simplified diagram of a standard gradient echo (GRE) pulse sequence.

The main pulses sequences used in this Thesis are briefly described below:

Gradient-recalled echo (GRE) employs an RF pulse, typically with a flip angle lower than 90°, and a series of gradients with varying polarities. This combination is designed to enforce coherence among the spins within the excited slice. As can be seen in Figure 2.2, after the RF pulse, there is a prephasing gradient lobe, which dephases the excited spins. Subsequently, these spins are rephased by the readout gradient lobe, resulting in the generation of an echo peak signal which presents a maximum at the TE. In this type of sequences the intensity of the echo peak will be primarily weighted by the T2* relaxation time, which depends on the tissue properties and also on the field inhomogeneities and susceptibility effects. On certain occasions, for example when reducing TR for obtaining T1w images, residual transverse magnetization can persist from cycle to cycle. To resolve this problem, in spoiled GRE sequences as spoiled gradient-recalled echo (SPGR), gradients and/or RF pulses (spoilers) are used to eliminate residual transverse magnetization.

Spin echo (SE) is composed of two RF pulses. The first one (typically of 90°) acts as an excitation pulse, and the second one (typically of 180°) acts as a refocusing pulse. The refocusing RF pulse introduces a phase inversion that corrects for accumulated phase variations resulting from chemical shifts, field inhomogeneities, and susceptibility effects. The echo peak materializes at the TE, which corresponds with twice the temporal interval between the two RF pulses. Thus, the refocusing RF pulse occurs at TE/2. A diagram of this sequence is shown in Figure 2.3. In this type of sequences the intensity of the echo peak can be weighted by the T1, T2 or PD, depending on the TE and TR selected for the acquisition (see Table 2.1).

Inversion recovery (IR) is the most common form of magnetization preparation. In this type of sequences, an 180° inversion pulse is used to flip the magnetization into the '-z'-axis. From this position the magnetization relaxes back to the steady

Table 2.1: Combinations of TE and TR values used in SE sequences to generate main contrast weightings. (Bernstein et al., 2004; Westbrook, 2015).

	TE [ms]	TR [ms]
T1-weighted	(10-30 ms)	(400-700 ms)
PD-weighted	(10-30 ms)	($\geq$ 2000 ms)
T2-weighted	($\geq$ 70 ms)	($\geq$ 2000 ms)

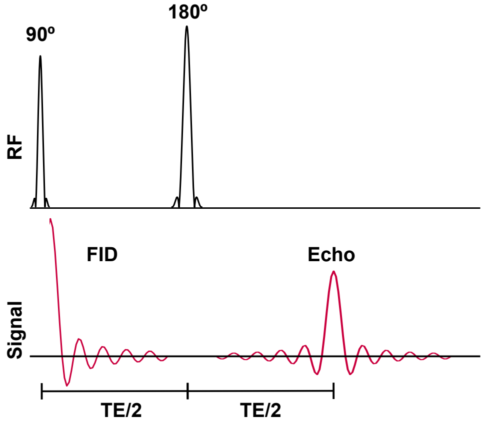


Figure 2.3: Simplified diagram of a standard spin echo (SE) pulse sequence.

state according to the specific T1 value of the tissue. Each tissue passes through zero at a different time, since each particular tissue presents an specific T1. Some time after the inversion pulse, the RF excitation is applied; time known as TI. During the RF excitation only tissues with non-zero longitudinal magnetization will produce an MR signal. Thus, the TI can be chosen in order to null the signal for a given tissue. This magnetization preparation mechanism is used, for example, in fluid attenuated inversion recovery (FLAIR) sequences.

2.1.4 Theoretical MR pulse sequences equations

Upon ideal conditions, the Bloch equation previously stated in Eq. 2.3 presents an analytical solution specific of the corresponding pulse sequence selected. Thus, for simple cases, the synthesis of MR weighted images can be performed with these well-known theoretical equations that describe MR intensity as a function of operator-selectable acquisition parameters and T1, T2, and PD parametric maps. In more complicated cases, more sophisticated methods are needed (Bittoun et al., 1984). In this Thesis, we synthesize weighted images corresponding to the following sequences: IR-GRE* (see Appendix B), spin echo (SE), gradient-recalled echo (GRE), and inversion recovery-spin echo (IR-SE), with respective equations Eq. (2.10)—(2.13):

---

*IR-GRE is equivalent to magnetization prepared rapid gradient echo (MPRAGE; Siemens), IR spoiled GRE (IR-SPGR or FSPGR with inversion activated or BRAVO; GE) and 3D turbo field echo (TFE; Philips) (Ellingson et al., 2015a)

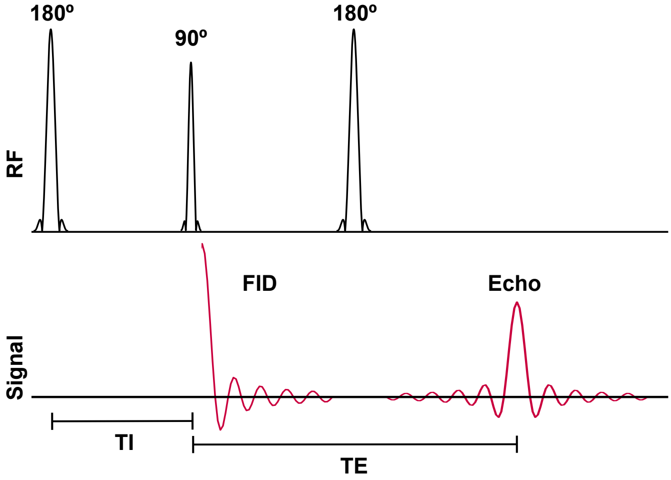


Figure 2.4: Simplified diagram of a inversion recovery (IR) spin echo (IR-SE) pulse sequence.

$$m_{\text{IR-GRE}}(\mathbf{x}) = \text{PD}(\mathbf{x}) \frac{1 - 2e^{-\text{TI}/T_1(\mathbf{x})} + e^{-\text{TR}/T_1(\mathbf{x})}}{1 + \cos(\alpha)e^{-\text{TR}/T_1(\mathbf{x})}} \sin(\alpha) e^{-\text{TE}/T_2(\mathbf{x})} \tag{2.10}$$

$$m_{\text{SE}}(\mathbf{x}) = \text{PD}(\mathbf{x}) \left[1 - 2e^{-(\text{TR}-\text{TE}/2)/T_1(\mathbf{x})} + e^{-\text{TR}/T_1(\mathbf{x})}\right] e^{-\text{TE}/T_2(\mathbf{x})} \tag{2.11}$$

$$m_{\text{GRE}}(\mathbf{x}) = \text{PD}(\mathbf{x}) \frac{1 - e^{-\text{TR}/T_1(\mathbf{x})}}{1 - \cos(\alpha)e^{-\text{TR}/T_1(\mathbf{x})}} \sin(\alpha) e^{-\text{TE}/T_2(\mathbf{x})} \tag{2.12}$$

$$\begin{aligned} m_{\text{IR-SE}}(\mathbf{x}) = \text{PD}(\mathbf{x})[1 - 2e^{-\text{TI}/T_1(\mathbf{x})} + 2e^{-(\text{TR}-\text{TE}/2)/T_1(\mathbf{x})} \\ - e^{-\text{TR}/T_1(\mathbf{x})}]e^{-\text{TE}/T_2(\mathbf{x})} \end{aligned} \tag{2.13}$$

with $\mathbf{x}$ the voxel location defined on some domain $\mathcal{X}$.

2.2 Post-contrast imaging

2.2.1 Contrast agents

Some transition metal elements with paramagnetic properties, such as gadolinium (Gd^{3+}) or manganese (Mn^{2+}), possess unpaired electron spins in their outer electron orbitals. The magnetic field generated by an electron is considerably stronger than that produced by a ^1H proton, making these paramagnetic elements excellent candidates for creating MRI contrast agents. These agents affect the T1 and T2 values of the tissues, thereby affecting tissue signal intensity and contrast.

Contrast enhancement in MRI can be achieved by either increasing or decreasing tissue signal intensity. The dominant contrast effect, whether it is T1 shortening (i.e., positive contrast agent) or T2 shortening (i.e., negative contrast agent), depends on several factors, including the dose of the contrast agent and the selected MRI pulse sequence.

A prime example of positive contrast agents are those based on Gd^{3+}. These agents reduce both T1 and T2 values of tissues, but due to the considerably longer T1 values compared to T2, their primary impact at low doses consists of T1 shortening. Consequently, tissues that accumulate such contrast agent appear bright or hyperintense in T1w images. This is why Gd^{3+} is known as a T1w-enhancement agent.

2.2.2 Tumor assessment with post-contrast imaging

The introduction of first GBCA in the market (Magnevist, Bayer Healthcare, Berlin, Germany) dates back to 1988 (Reeder, 2014). Since then, there has been a quantum leap in the usage of GBCAs in radiology with more than 30 million worldwide patient administrations per year at the time being (Lohrke et al., 2016). Specifically, GBCAs have demonstrated their exceptional utility in visualizing the CNS because of their ability to pass through disruptions in the BBB. The BBB is the physical interface between the CNS and systemic circulation; it tightly regulates what enters and is removed from the brain parenchyma, being fundamental in maintaining brain homeostasis. It is known that the BBB is disrupted during tumor progression and, after that, it is referred to as the blood tumor barrier (BTB) (Arvanitis et al., 2020). One important feature of the BTB is that is more permeable than the BBB, as can be seen in Figure 2.1.

The standard procedure used in clinical practice begins with the acquisition of a baseline pre-contrast T1w image to establish the initial state. Subsequently, a GBCA is administered through a bolus intravenous injection. The standard recommended dose depends on the specific GBCA, but for the majority of them it is 0.1mmol/kg (Reeder, 2014). After a waiting period of approximately 5-10 minutes, another T1-weighted acquisition is performed to detect any potential leakage of the GBCA into the brain parenchyma, which manifests as contrast enhancement (Warntjes et al., 2018). The second image is referred to as post-T1w. During diagnosis, both images (T1w and post-T1w) are usually shown side by side, and the evaluation of T1w-enhancement is typically based on a subjective visual comparison between them.

While contrast enhancement assessment plays a pivotal role in the evaluation of high-grade gliomas, it is worth noting that these tumors are also recognized for their infiltration into the surrounding peritumoral edema. However, detecting tumor infiltration is challenging when relying solely on visual inspection of conventional weighted images (Blystad et al., 2020, 2017; Müller et al., 2017). The assessment of treatment effects and tumor evolution might be also difficult using only pre- and post-contrast weighted images. Some treatments can induce changes within a tumor, which does not necessarily indicate real reduction of tumor burden (Ellingson et al., 2015b; Lescher et al., 2015). Therefore, quantitative measurements could provide valuable additional information in such cases.

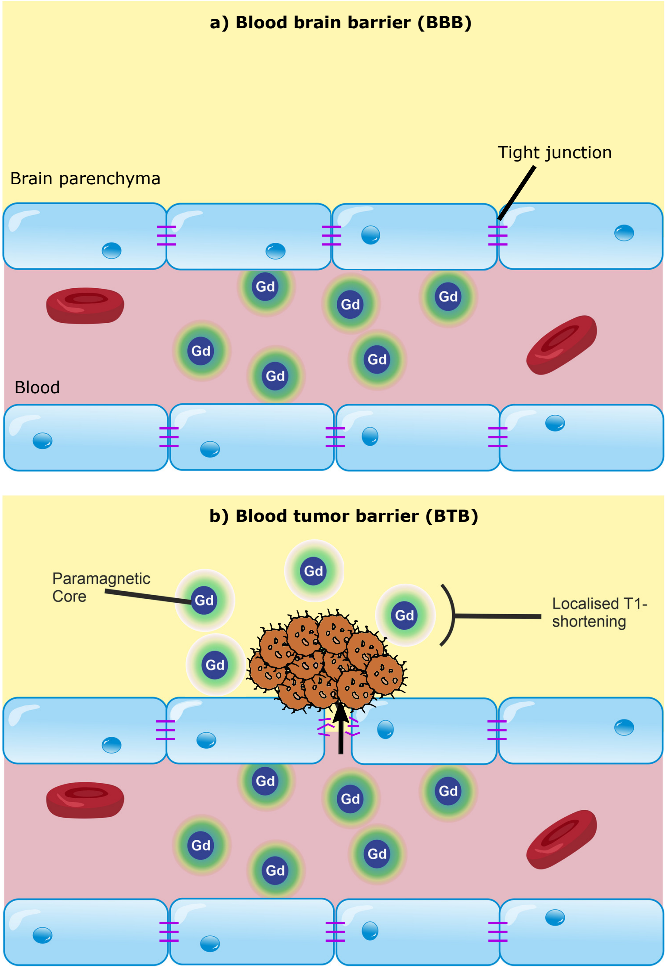


Figure 2.5: Schematic representation of the differences between (a) the blood brain barrier (BBB) and (b) the blood tumor barrier (BTB). The BTB is characterized for being more permeable than the BBB, allowing the Gd to pass through.

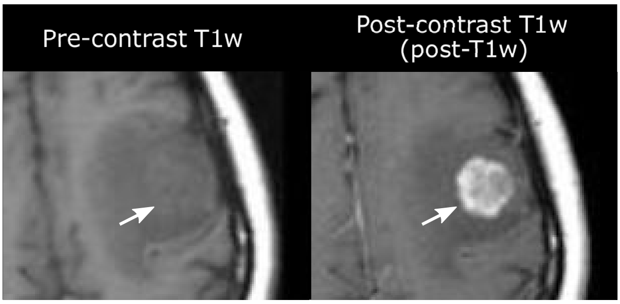


Figure 2.6: Well-defined T1w signal enhancement visible within a tumor in the left frontal lobe in the image acquired after a contrast agent administration.

2.2.3 Limitations and Pitfalls

The MR acquisitions performed after an injection of a GBCA present some disadvantages compared with those that do not use such agents. These disadvantages were introduced in Section 1.1, and are thoroughly described next:

- **Safety concerns**: although GBCAs are generally deemed safe, between 0.7 to 2.4% of injected patients suffer from mild adverse reactions and a lower rate from severe complications (Forghani, 2016). Some studies found that the exposure to GBCAs can potentially initiate the onset of nephrogenic systemic fibrosis (NSF), a rare but life-threatening complication that may occur under specific conditions (Fraum et al., 2017; Grobner, 2006). Additionally, safety concerns have recently arisen due to the possible deposition of the GBCAs in the body and the brain (Gulani et al., 2017), especially, in patients who need to undergo several follow-up acquisitions, as it is the case of oncological patients. Hence, some studies have found signal hyperintensity on non-contrast-enhanced T1w images (involving the dentate nucleus and globus pallidus) in patients that have previously been administered multiple doses of GBCAs (Errante et al., 2014; Kanda et al., 2014; Runge, 2016).
- **Environmental issues**: GBCAs have come out as water pollutants. Hence, recent studies have found presence of GBCAs in wastewater (Inoue et al., 2020; Rogowska et al., 2018).
- **Scan time**: post-contrast acquisitions require longer scan time, which reduce the accessibility to MR scans and might also lead to motion artifacts and, consequently, to extra efforts in re-acquiring or post-processing (Xie et al., 2022).
- **Discomfort**: the intravenous injection of the GBCA, as well as the prolonged scan time, lead to patient discomfort.
- **Cost**: the usage of GBCAs increases the need of skilled manpower, hardware and thus, costs (Shankar et al., 2018).

Safety and environmental concerns of GBCAs could be tackled by the usage of new contrast materials (Wesolowski and Kaiser, 2016) or novel sequences in which the injection is not required, such as amide proton transfer (APT) imaging (Zhou et al., 2003) or arterial spin labeling (ASL) (Petersen et al., 2006). Nevertheless, these techniques, albeit promising, still require complex and expensive acquisition schemes. In addition, the resulting images are usually more difficult to analyze.

2.3 Quantitative MRI

2.3.1 Clinical applications of parametric maps for tumor diagnosis

As previously described in Subsection 2.1.1, T1 and T2 values depend on both the strength of the B_0 magnetic field and the tissue itself. In addition, the T1 and T2 values might also be different in healthy and pathological tissues. Thus, parametric maps could act as biomarkers of several diseases, such as epilepsy (Conlon et al., 1988), stroke (DeWitt et al., 1987), multiple sclerosis (Larsson et al., 1989), and brain tumors (Badve et al., 2017; Yankeelov et al., 2011), among others.

Particularly, for tumor diagnosis, potential applications of T1 and T2 maps are tumor grading and characterization. Several studies have demonstrated longer T1 and T2 values within tumors compared to normal white matter (Kurki and Komu, 1995; Nunez-Gonzalez et al., 2022; Pirkl et al., 2021). Badve et al. (2017) reported differences in T1 values between peritumoral regions of low-grade gliomas and glioblastomas. As for T2 values, early research indicated that meningiomas exhibit shorter T2 values compared to astrocytomas (Kjær et al., 1991) and gliomas (Naruse et al., 1986; Oh et al., 2005). Additionally, it was observed that glioblastomas exhibit a single T2 component when predominantly composed of solid tissue, but they display two T2 components when consisting of a solid/necrotic mixture (Naruse et al., 1986).

Another potential application is monitoring tumor progression, which is a key task in glioblastoma patient management. Longitudinal T1 mapping (i.e., acquisition of T1 maps at several time points) has shown promising results offering more precise quantitative insights into tumor dynamics, especially with regard to tumor enhancement. Hattingen et al. (2017) have shown in a comparison of quantitative T1 mapping before and after the GBCA injection that areas with T1 values > 2051 milliseconds at 3T could predict the presence of enhancing brain tissue. T1 mapping also offers the advantage of detecting tumor enhancement at an earlier stage, and more extensively, compared to conventional post-T1w images (Lescher et al., 2015; Müller et al., 2017).

Quantitative T1 mapping could also be a valuable tool to distinguish between tumor recurrence and radiation necrosis, a common challenge encountered in gamma knife radiosurgery. In a study involving patients with brain metastases who subsequently underwent biopsy, T1 mapping was conducted at two time points: 5 minutes and 60 minutes after the administration of a contrast agent (Wang et al., 2018). This research revealed that the most effective parameter for diagnosis was the difference in T1 values between time points. This finding suggests the eventual clinical usefulness of this approach for distinguishing between tumor growth and radiation-induced brain changes.

2.3.2 MRI Relaxometry: T1 and T2 mapping

In the classical methods to estimate parametric maps, a set of N weighted images (i.e., signal measurements) is customarily acquired with the same pulse sequence but with a varying acquisition parameter. Then, the parametric map estimation problem consists on finding the parameter values which best fit the measurements of the relaxation signal model. Thus, the map is computed by voxel-wise fitting a known relaxation model (which is dependent on the pulse sequence) to the intensity of the images. Although there are different ways to perform T1 and T2 mappings, next we describe the most popular techniques.

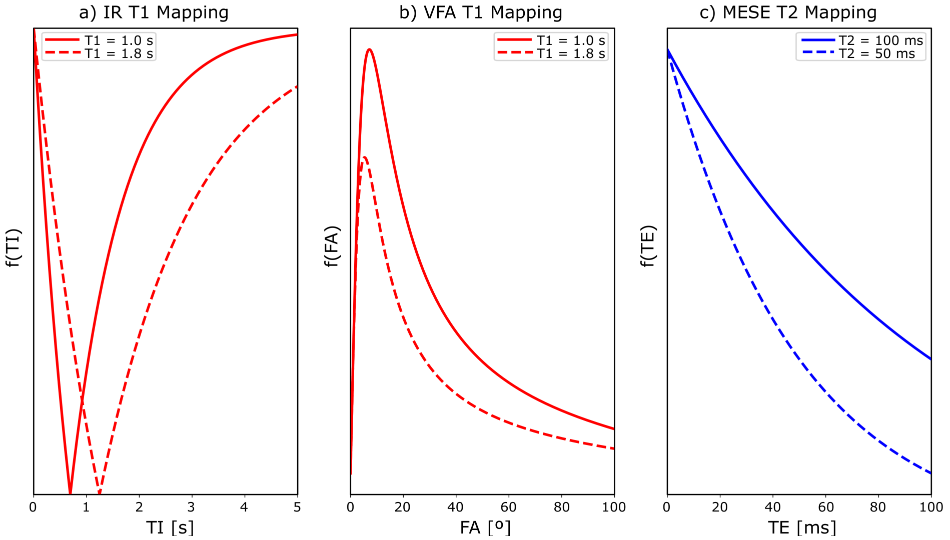


Figure 2.7: Signal models employed in different relaxometry techniques for T1 and T2 mapping. a) Signal model of a IR sequence as a function of TI for two different T1 values, b) Signal model of a VFA sequence as a function on α for two different T1 values. c) Signal model of a ME-SE sequence as a function of TE for two different T2 values.

T1 mapping. Two sequences are usually employed for T1 mapping, namely, inversion recovery (IR) and VFA. The former, which is considered the gold-standard, requires the acquisition of a set of IR MR images with different TIs. The signal model of an IR sequence for varying TI is shown in Figure 2.7a). The latter consists of the acquisition of a range of steady-state SPGR MR images over a set of flip angles (see Figure 2.7b) for a representation of a VFA signal modal for varying α). The TR can be much shorter in the latter than in the former, therefore with VFA higher resolution T1 maps can be obtained with a reduced acquisition time.

T2 mapping. One of the most employed sequences for T2 mapping is the multi echo-spin echo (ME-SE). In this sequence, within one acquisition, multiple echoes are generated by applying a train of refocusing RF pulses (typically 180°) after a 90° RF excitation. The TE vary from one echo to another, and different images are acquired at each TE, which allows the estimation of the T2 map. The signal model of an ME-SE sequence for varying TE is shown in Figure 2.8a).

2.3.3 Challenges in MRI Relaxometry

Several factors might affect the accuracy of the estimated parametric maps and the performance of the relaxometry methods. Also, others limit their utility in the clinical practice. Next, we summarize some of these sources of errors and limitations.

- **Scan time**. The total scan time of an MR relaxometry sequence is proportional to the number of weighted images acquired. To be able to perform accurate estimations of the parametric maps, a sufficiently large set of weighted images needs to be acquired, which imply a time-consuming acquisition.

- **Motion**. Long relaxometry acquisitions make them prone to motion artifacts. These type of artifacts can appear in an image due to patient movement during the k-space readout. In addition, if the patient moves from the acquisition of one weighted image to another, the acquired weighted images could be misaligned, leading to inaccuracies in the estimation of the parametric maps.

- **Fitting procedures**. The usage of closed-form formulas to approximate the signal models has consequences in the accuracy of the estimation. In addition, some commonly used fitting procedures implicitly assume some reality simplifications, which might also affect the estimated values of the parametric maps.

- **Flip angle inhomogeneities**. Precise knowledge of the flip angle is important for the accurate estimation of the T1 and T2 maps. Nevertheless, the flip angle chosen within the scanner may not necessarily match the ones that are actually transmitted for each position of the scanned object.

- **Independent parametric mapping procedures**. Most of the proposed relaxometry methods only provide information of a single parameter at a time (Barbieri et al., 2020; Clare and Jezzard, 2001; Ramos-Llordén et al., 2018), so additional sequences are needed to obtain the different parametric maps.

2.4 Synthetic MRI versus MR image translation

Synthetic MRI is a technique that computes quantitative parametric maps from some acquired weighted images and then synthesizes new MR weighted images from those estimated parametric maps. The full pipeline is based on a three-step procedure that was previously introduced in Section 1.1 and it is outlined in Figure 2.8. This approach is different from conventional weighted MRI, which is based on the direct acquisition of several types of weighted images, but with no quantification.

A typical Synthetic MRI approach involves the computation of T1, T2, and PD maps, which can then be utilized to synthesize customized weighted images with specific acquisition parameters (such as TE, TR, TI, etc.) through the use of a signal model. This ability to synthesize images with specific parameters sets Synthetic MRI apart from quantitative MRI. While quantitative MRI serves as a fundamental component of synthetic MRI, its primary emphasis is on estimating parametric maps, rather than on synthesizing weighted images. Nevertheless, the process of synthesizing a weighted image from quantitative maps using a signal model is relatively straightforward. It is worth noting that weighted images hold particular significance in clinical practice, as they constitute the primary resources for routine radiological assessments.

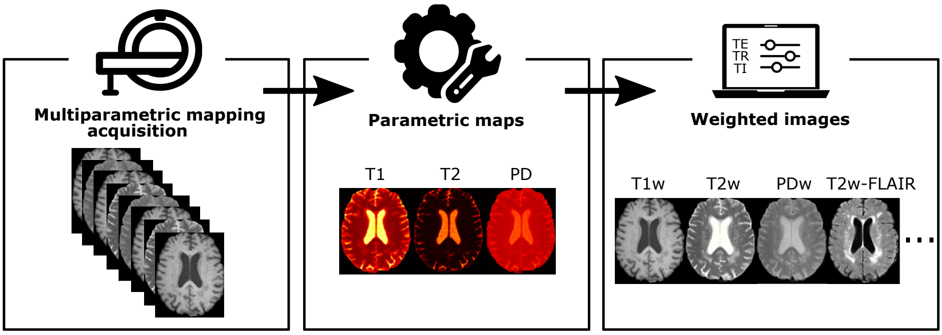


Figure 2.8: Pipeline of a three-step typical Synthetic MRI approach. First, a multiparametric mapping acquisition is performed in the scanner. This acquisition allows the estimation of the T1, T2, and PD parametric maps in the second step. Finally, in the third step, an arbitrary number of weighted images with different sequence parameters are synthesized from the parametric maps.

2.4.1 Related work

Different methods that fit the conventional three-step definition of Synthetic MRI (Bobman et al., 1985) have been proposed. Gulani et al. (2004) proposed an steady-state precession (IR-TrueFISP) sequence in which a series of different IR time-delayed TrueFISP images are acquired to quantify the parametric maps. Then T1w, T2w, PDw, and T2w-FLAIR images were synthesized from these maps. Warntjes et al. (2008) proposed a multiecho acquisition of a saturation-recovery turbo spin-echo readout (QRAPMASTER)† for the quantification of T1, T2, PD, and B1 inhomogeneity parametric maps. After quantification, T1w, T2w, and T2w-FLAIR images were synthesized (Blystad et al., 2012). Finally, Cheng et al. (2020) suggested a multipathway multiecho (MPME) sequence using an unbalanced steady-state sequence with two different flip angles and resolution scans to quantify T1, T2, T2*, and B0 and B1 inhomogeneities parametric maps. Then, the authors showed the synthesis of T1w, T2w, PDw, and T2w-FLAIR, and magnetization prepared rapid gradient echo (MPRAGE) images using a neural network. Nevertheless, this latter method suffers from noise amplification due to the multiple processing steps which lead to somewhat noisy maps and synthesized images.

It is important to note, however, that all of these methods require very specific sequences or private protocols scarcely available in clinical scanners. Also, these quantitative sequences are focused on obtaining the parametric maps, but they are not commonly used for diagnosis purposes in the clinical routine at the time being.

In contrast, MR image translation approaches (Chartsias et al., 2017; Dar et al., 2019; Sohail et al., 2019) share the same objective as step 3) of Synthetic MRI

---

† QRAPMASTER is nowadays referred to as a multidynamic multiecho (MDME) sequence (Tanenbaum et al., 2017).

definition, although their input is one or more weighted images as opposed to parametric maps. Hence, they are not flexible as to which modalities can be generated, since most of them are tailored for a specific application where, given some input image modalities, new predefined image modalities are synthesized. For example, in Dar et al. (2019) the authors synthesize T1w from T2w images. These methodologies limit themselves to the image modalities used in the learning stage since the potential of parametric maps to synthesize any weighted image is not employed.

2.5 Deep learning: From mathematical models to MR image synthesis

2.5.1 Basics of neural networks

A neural network (NN) is comprised of interconnected layers of basic units called nodes or artificial neurons that perform simple mathematical operations to detect patterns in the input data. Thus, a NN is typically composed of an input layer which receives data represented by a numeric value, multiple hidden layers which perform the computations and, finally, an output layer which predicts the output.

The basic building block of a NN is the aforementioned artificial neuron. McCulloch and Pitts (1943) were pioneers on describing the idea behind the artificial neuron. This notion was designed to mimic the way a brain neuron was thought to work. To this end, they modeled a simple NN using electrical circuits. As can be seen in Figure 2.9, an artificial neuron takes one or more input values (x_1, x_2, ... x_n), multiplies each of them by their corresponding weights (w_1, w_2, ... w_n) and then adds them up together with a bias (b). Afterwards, the addition z is passed through an activation function ($g(.)$) to get the predicted output ($\hat{y}$).

The activation function is a mathematical function ($g(.)$) used to induce nonlinearity into the output of a neuron. This way, according to Cybenko (1989), a NN with only one hidden layer is capable of always approximating a multi-variant continuous function. Different activation functions, such as *sigmoid*, hyperbolic tangent, rectified liner unit (ReLU) (Nair and Hinton, 2010), etc., can be employed depending on the characteristics of each problem. Specifically, ReLU activation is defined as $g(z) = \max(0, z)$ and, therefore, the output is always non-negative. This activation is characterized for being computationally efficient and not suffering from the vanishing gradient problem. In contrast, its main issue is known as "dying ReLU" and is caused by a neuron with a negative value which may never activate. To avoid this issue, the Leaky ReLU activation function is defined as $g(z) = max(az, z)$, where a is a small constant. The Leaky ReLU activation is equivalent to ReLU activation when z is positive. However, the Leaky ReLU activation returns a small negative value proportional to z when z is negative.

Thus, the outcome of the activation function in turn decides if a neuron is activated or not. The output of an activated neuron is passed to other neurons until the last layer of the network is reached. This process is known as *forward propagation*. The neuron in the output layer then projects the result.

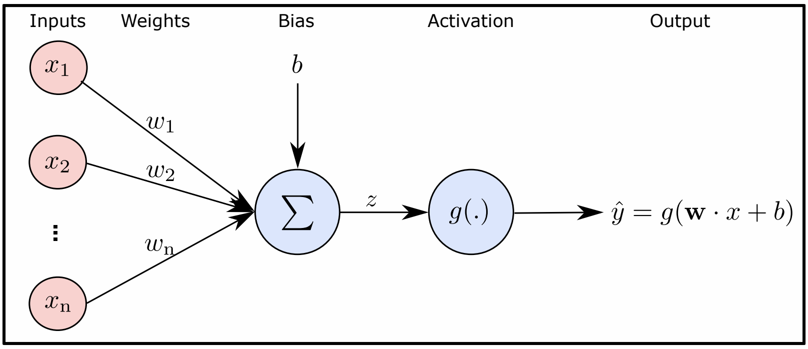


Figure 2.9: Schematic representation of an artificial neuron or node with a total of n input values.

2.5.2 Types of learning

Training of a NN consists in an iterative optimization process which involves adjusting the model (i.e., the NN weights) to generate accurate predictions from the input data. One of the most common methods by means of which the network is trained is known as *Backpropagation*, which is a type of gradient descent algorithm. This method refers to the calculation of a gradient (i.e., the derivative of the loss function) on each weight in the NN for each training element. The gradient of each weight indicates how to re-adjust that corresponding weight to minimize the loss function, thereby resulting in a more accurate output. Different types of NN learning techniques can be distinguished according to the characteristics of the training process.

- **Supervised learning:** this learning algorithm is employed when there exist labeled data (i.e., the association between the input data and the corresponding correct outputs). This type of learning enables the optimization algorithm to establish a mapping between input and output. The objective of supervised learning is to obtain a model capable to generalize and perform accurate predictions for new, unseen input data.

- **Unsupervised learning:** this learning algorithm is employed to identify patterns in unlabeled data. This approach is typically employed when there is a scarcity of training data or when the data is too intricate for labeling. Examples of unsupervised learning algorithms are clustering, which is useful for organizing or finding patterns in data, or association rule mining, which allows to find relationships between data items.

- **Self-Supervised learning:** this learning algorithm, which is a variation of unsupervised learning, enables the derivation of knowledge directly from the input data, eliminating the need of explicit labeling. This learning algorithm is employed, for example, in physics-informed NNs. These type of NNs, which have recently gain popularity, are able to create more powerful models by means of the incorporation of physical principles and prior scientific knowledge into the training process.

- **Reinforcement learning:** this learning algorithm is neither based on supervised nor unsupervised learning. In contrast, in a reinforcement learning algorithm, the learning agent goes from one state to another and receives rewards only on success but not on failure. This way, the agent learns from the environment.

In this Thesis we employ both supervised and self-supervised learning techniques.

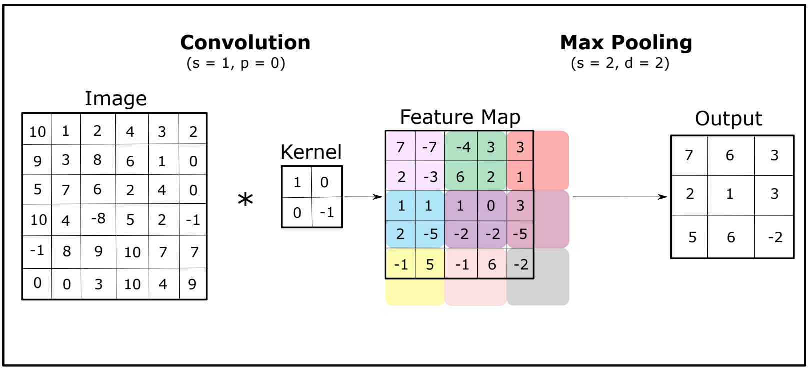
2.5.3 Convolutional neural networks

A CNN is a class of NN that is specialized in processing data with a grid-like topology, such as an image. In this Thesis we focus on this type of NN. The main layers in CNNs are convolutional, pooling, and dense layers. The characteristics of these types of layers are described next.

Convolutional layers are the core building block of the CNN. The main goal for a convolutional layer is to detect features in the input image. Thus, convolution leverages sparse interaction, parameter sharing, and equivariant representation. This layer performs a dot product between the filter or kernel and the specific area of the image which is being scanned by the kernel. The kernel is a matrix with a square shape whose values correspond with the set of learnable parameters (i.e., weights). Hyper-parameters of this type of layers include the size of the squared-shape filter (w) and also the depth or number of filters, which correspond with the number of features extracted in the convolution operation. During the forward pass, the kernel slides across the height and width of the image. This produces a two-dimensional representation of the image known as an activation map that gives the response of the kernel at each spatial position of the image. Other hyper-parameter are the sliding size of the kernel so-called stride (s) and the padding (p), which involves adding extra pixels around the border of the input before the convolution in order to preserve spatial dimensions.

Pooling layers decrease the spatial size of the representation by performing a downsampling operation. These layers help in reducing the number of required weights and, thus, the computational load. The pooling operation is processed on every slice of the representation individually. There are several pooling operations, but the most popular is max-pooling, in which the output corresponds with the maximum value of the neighborhood (see Figure 2.10 for a simple example of this layer). Two hyper-parameters characterize this type of layer: spatial extent (d) and stride (s); the first one specifies the size of the boxes that will be scaled down, and the second one refers to the sliding step. Note that max-pooling layers have no weights, so training does not affect them.

Dense layers are also known as fully-connected layers, since the neurons in these layers have full connectivity with all neurons of the preceding and following layers. The output of these layers can be computed as we described above as the addition of weighted inputs and a bias effect. The main hyper-parameters of dense layers are the number of neurons and the activation function to be used.



Image

10	1	2	4	3	2
9	3	8	6	1	0
5	7	6	2	4	0
10	4	-8	5	2	-1
-1	8	9	10	7	7
0	0	3	10	4	9

Kernel

1	0
0	-1

Feature Map

7	-7	-4	3	3
2	-3	6	2	1
1	1	1	0	3
2	-5	-2	-2	-5
-1	5	-1	6	-2

Output

7	6	3
2	1	3
5	6	-2

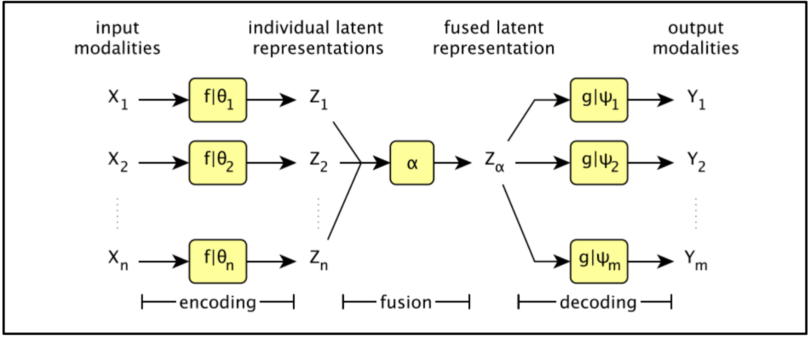
Figure 2.10: Example of the concatenation of a convolutional layer with a kernel size (w = 2), stride (s = 1), and no padding, followed by a max-pooling layer with spatial extent (d = 2) and stride (s = 2).

2.5.4 Encoder-decoder CNN architecture for medical image translation

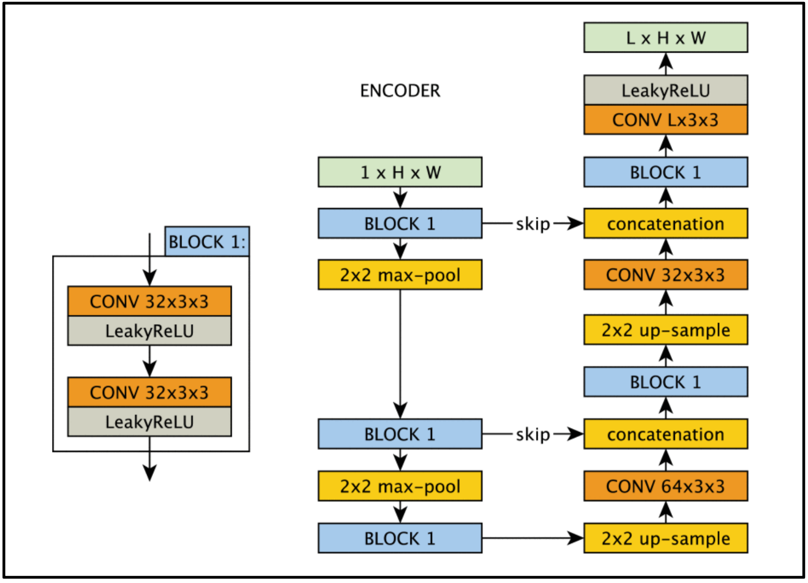
Chartsias et al. (2017) proposed a DL method for translation between different MR image modalities based on a fully convolutional neural network (FCNN). Thus, this network takes as input 2D slices of any subset of its inputs, and synthesizes the corresponding 2D slices in all output modalities. As can be seen in Figure 2.11.a), the network is composed of three stages, namely, encoding, fusion, an decoding. The model is trained end-to-end with gradient descent minimizing a multi-component loss function, which balances between modality-invariance and the retention of modality specific information.

1. **Encoding stage.** The purpose of the encoders is to embed input images modalities into multi-channel latent representations. Thus, one encoder is included for each input modality. The architecture of each encoder is inspired by the well-established U-Net architecture (Ronneberger et al., 2015) as shown in Figure 2.11.b). Specifically, the network includes two downsample (and upsample) steps. The concept behind the architecture of U-Nets, which involves down-sampling followed by up-sampling and incorporates skip connections, is to enable the network to leverage information from broader spatial scales than the filters themselves, all while retaining valuable local information. Furthermore, the inclusion of skip connections aids in maintaining a smooth gradient flow during the training process. The authors employed Leaky ReLU (Maas et al., 2013) to facilitate the training of the network and to improve the quality of the latent representation. They also used a stride of 1, and padding of the images by repeating the border pixels so that the final output has the same width and height as the original input.
2. **Fusion stage.** The goal of this stage is to combine the individual latent representations produced by the encoders into a single fused representation. The authors pursue not only to preserve commonly represented features, but also to retain unique features expressed in one modality but not the others.

(a) CNN overview



(b) Encoder architecture



(c) Decoder architecture

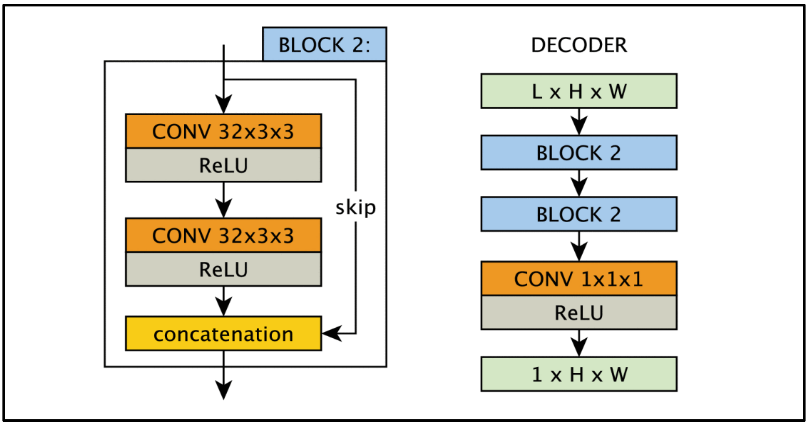


Figure 2.11: Encoder-Decoder architecture for image translation. (a) Overview of the FCNN. (b) Encoder architecture inspired by the U-Net (Ronneberger et al., 2015). (c) Decoder architecture. (Figures take from Chartsias et al. (2017))

To this end, they use the pixel-wise max function to combine individual latent representations of each encoder into a fused latent representation. The use of the max means that, in each channel, each pixel of the latent representation has exactly the value of the corresponding pixel in one of the original latent representations. In particular, if the signal is large and positive in one constituent latent representation, then it will be chosen for the fused representations. Thus, the size of the fused latent representation is exactly the same as the individuals counterparts. Note that the use of the max operation does not bias the method towards bright final outputs, as the intensities of the synthesized images depend on the decoding step.

3. **Decoding stage.** The decoders are in charge of mapping the multi-channel fused latent representation into the desired output modalities. Thus, one decoder is included for each output modality. The architecture of each decoder is a FCNN as shown in Figure 2.11.c). The authors kept the decoder shallower than the encoder to encourage the latent representation to contain the useful information in a simple way.

Throughout the methods proposed in this Thesis, we use different variations of this encoder-decoder CNN architecture to perform the T1, T2 and PD multiparametric mapping from conventional weighted images. The number of inputs and outputs were modified accordingly. Also, the learning approach was adapted depending on the labeled datasets available for training. When no labeled dataset was available, self-supervised learning approaches where employed by means of including physical priors in the computation of the loss function.

2.5.5 Advanced NN architectures for image synthesis

Recently, sophisticated NN architectures have been developed. These architectures have shown impressive results in an ample variety of image processing tasks, and particularly, in medical images synthesis. Although, these network architectures have not been used in this Thesis, we consider them relevant for future work and, consequently, their main characteristics are briefly introduced next:

Generative Adversarial Networks A generative adversarial network (GAN) is a type of NN initially proposed by Goodfellow et al. (2014). It is composed of two networks, namely, a generator and a discriminator. These two networks undergo a joint training in a two-player minimax game framework, where the generator aims to deceive the discriminator by producing outputs with real appearance, while the discriminator tries to distinguish whether its inputs are real or generated. Equilibrium is attained when the discriminator can no longer differentiate between real and generated inputs.

Thus, for the particular case of image translation, GANs have been proposed to learn the distribution of target image modalities conditioned on source image modalities, thereby, capturing higher frequency details compared to traditional CNNs. Specifically, GANs methods have been employed in several medical image translation applications including MRI to CT (Liu et al., 2021), low to high resolution (Wang et al., 2020; You et al., 2022), 3T to 7T (Li et al., 2023), and multi-contrast weighted MRI synthesis (Dar et al., 2019; Sohail et al., 2019), among others.

This network architecture, however, also presents some limitations. On the one hand, some major problems related to its training are non-convergence, collapse mode and vanishing gradients. On the other hand, the generators of the GANs are normally based on CNNs, which present, in general, difficulties to generalize to abnormal anatomies for different subjects.

Vision Transformers A vision transformer (ViT) is a type of NN based on the transformer architecture (Vaswani et al., 2017) but with some modifications that make them more suitable for image processing tasks. Thus, ViTs represent images as a sequence of small rectangular regions called patches. After splitting the image into patches, each patch is encoded as a vector, capturing its distinctive features. These features are commonly extracted using a CNN. Subsequently, the vectors are input into a Transformer encoder, which consists in a stack of self-attention layers. Self-attention is a mechanism that enable the model to weigh the importance of different elements in a sequence and, thus, to learn global dependencies among the patches. Thus, the potential of the ViTs resides in their ability to learn global features of images due to the possible attention to any part of the image, regardless of its location. In contrast, their main problems are that they are not as interpretable as CNNs and their computationally expensive training.

Recently, Dalmaz et al. (2022) have proposed the first adversarial model for medical image synthesis with a generator based on a ViT. The method is employed to translate between multi-contrast MR weighted images and, also, for MRI to CT translation. After that, other methods based on Transformers have been proposed for MR weighted image translation (Liu et al., 2023).

Part II

Synthetic MRI from routine sequences

3

A deep learning approach for Synthetic MRI based on two routine sequences and training with synthetic data

Contents

The work in this chapter has been published in:

- **Moya-Sáez, E.**, Peña-Nogales, Ó., de Luis-Garcia, R., and Alberola-López, C. A deep learning approach for synthetic MRI based on two routine sequences and training with synthetic data. *Computer Methods and Programs in Biomedicine.* 2021; 210, 106371.
- **Moya-Sáez, E.**, Peña-Nogales, Ó., Sanz-Estébanez, S., de Luis-Garcia, R., and Alberola-López, C. CNN-based synthesis of T1, T2 and PD parametric maps of the brain with a minimal input feeding. *ISMRM & SMRT Virtual Conference & Exhibition*, Virtual, August 2020; 3806.

3.1 Purpose

Synthetic MRI has received significant attention during the past decade due to its ability to generate realistic MR images and, consequently, to reduce acquisition time and/or to retrospectively enhance patient throughput. This technique implies the computation of T1, T2, and PD parametric maps, which, in turn enable the synthesis of other weighted images with realistic resemblance. As we previously introduced in Chapter 2, the computation of parametric maps can be performed by means of classical relaxometry sequences or fast multiparametric techniques. The former require, in general, long acquisition time, whereas the latter require scarcely available software.

Alternatively, DL could be a potential solution for the computation of parametric maps from conventional, and therefore, widely available weighted images. Hence, parametric maps, could be easily computed in both preexisting databases and in newly acquired datasets without additional acquisition time. However, the lack of large, public datasets with both weighted images and the corresponding parametric maps could be one of the main limitations that hinders the usage of these DL techniques.

In this chapter we propose a joint Synthetic MRI approach based on DL for the computation of the T1, T2, and PD parametric maps from only a pair of inputs, namely, a T1w and a T2w. The input images are acquired with conventional sequences widely used in clinical routine instead of specific multiparametric sequences. After that, a number of weighted images unseen by the network can be satisfactorily synthesized out of the parametric maps computed with DL. In addition, a new training strategy based on a synthetic dataset is proposed. This way, we overcome the lack of large datasets with quantitative parametric maps.

3.2 Methods

The DL approach proposed in this work for the computation of the parametric maps uses an adaptation of the CNN proposed by Chartsias et al. (2017), which has been previously described in Section 2.5. We propose to train this network in a supervised way with a synthetic dataset generated by the synthesis of the T1w and the T2w input images from their corresponding parametric maps. An overview of the proposed dataset generation, CNN training approach, and validation procedures are shown in Figure 3.1.

3.2.1 Synthetic dataset generation

We create a synthetic dataset with 120 brain volumes starting from the anatomical model of a normal brain obtained with BrainWeb (Cocosco et al., 1997). The pipeline to create the synthetic dataset is described in the next four steps (see Figure 3.1.a).

First, we created 120 different sets of T1, T2, and PD maps from the BrainWeb anatomical model by giving uniformly distributed random values to the white

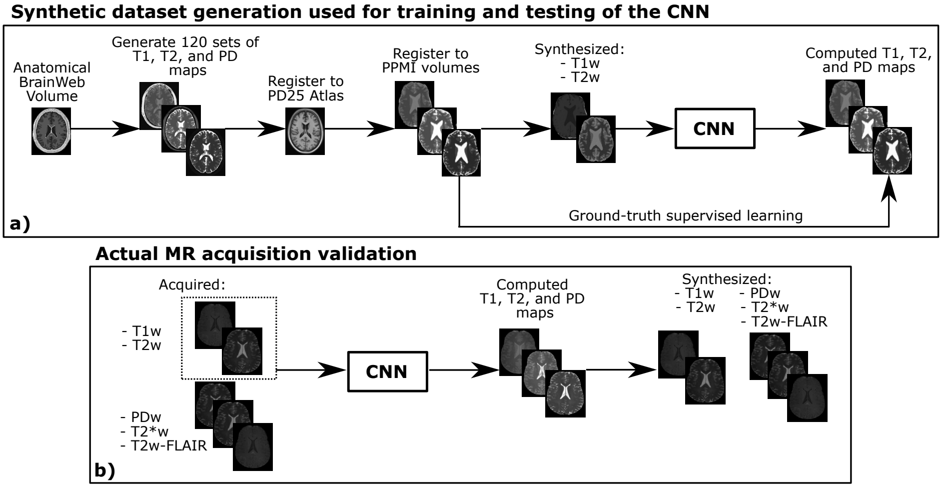


Figure 3.1: Pipeline of the proposed training and validation approaches. a) Synthetic dataset generation used for training and testing of the supervised CNN. b) Validation of the CNN with actual MR brain acquisitions.

matter (WM), grey matter (GM), and cerebrospinal fluid (CSF) of each map, one value for each label. The exact ranges of T1, T2, and PD values defined for each parameter in each tissue are shown in Table 3.1. Note that these ranges are within those reported in the literature for 3T MR scanners (Bojorquez et al., 2017). Also, additive Gaussian noise was added to each volume with distribution $\mathcal{N}(\mu = 0, \sigma = 0.01)$.

Table 3.1: Ranges of the T1, T2, and PD values for each of the three considered tissues in the anatomical brain model (WM: white matter, GM: grey matter, and CSF: cerebrospinal fluid) for the generation of 120 synthetic brain volumes. The specific value for each parameter in each tissue and volume is selected from a uniform distribution within these ranges. T1 and T2 values are given in seconds.

	T1 [s]	**T2** [s]	**PD**
WM	0.80 - 1.10	0.055 - 0.075	0.65 - 0.72
GM	1.40 - 1.60	0.075 - 0.120	0.77 - 0.82
CSF	4.50 - 4.80	1.20 - 1.60	1.20 - 1.30

Second, in order to introduce spatial variability across brain volumes (i.e., brains with different anatomical features), the maps were affine-registered to the PD25 atlas (Xiao et al., 2017) using the FLIRT tool of FSL (Jenkinson et al., 2012). Then, each set of maps was non-linearly registered to one out of the 120 different T1w volumes selected from the PPMI database (www.ppmi-info.org), with the FNIRT tool of FSL as described in Peña-Nogales et al. (2019). The size of each of these maps is of 240 x 176 with 256 slices. Subsequently, all sets of parametric maps were skull-stripped.

Third, for each set of skull-stripped T1, T2, and PD maps, a pair of weighted images was analytically synthesized. A T1w image was synthesized as a IR-GRE acquisition (see Eq. (2.10)) with TE = 3 ms, TR = 6.44 ms, TI = 900 ms, and $\alpha = 10°$. A T2w image was synthesized as a SE acquisition (see Eq. (2.11)) with TE = 85 ms and TR = 4000 ms. These particular sequences and parameter sets were chosen to match the actual MR brain acquisitions described in Table A.2 of Appendix A. Note that the weighted images have the same dimensions as the parametric maps (i.e., 240 x 176 with 256 slices).

Finally, we normalized the T1w and T2w images by dividing each of them by its average intensity without considering the background. This facilitates convergence of the CNN during training without altering image properties due to their qualitative nature.

3.2.2 Network training with the synthetic dataset

The aforementioned synthetic dataset was used to train an adapted version of the CNN described in Chartsias et al. (2017); our adaptation pursued to perform an end-to-end mapping function to transform the input T1w and T2w images to their corresponding set of T1, T2, and PD parametric maps (see Figure 3.1.a). Specifically, the weighted images were input to two encoders which embed these inputs into multi-channel latent spaces with the same image size as the inputs. Note that the CNN processes the inputs as 2D slices. The number of channels used is 16. Then, the latent representations of the input are fused into a single 16-channel representation using a maximum pixel-wise function between each pair of corresponding channels. This fused latent representation is next input to three decoders to obtain the three desired parametric maps.

Supervised training was carried out using the cost function proposed in Chartsias et al. (2017). This cost function minimizes 1) the mean absolute error (MAE) between the ground-truth parametric maps and the output's decoders (i.e. the synthesized parametric maps), and 2) the mean pixel-wise variance between latent representations. The model was trained through a mini-batch approach with a batch size of 8 images using Adam optimizer (Kingma and Ba, 2014) with a learning rate of 1×10^{-5}. We performed the training with early-stopping to avoid overfitting. From the 120 brain volumes of the synthetic dataset with a set of three parametric maps and two weighted images each, we used 70 for training (17920 slices), 36 for the early-stopping validation (9216 slices), and 14 for test (3584 slices).

The adapted CNN is coded in Python with Keras. We ran the code using the TensorFlow backend on a single NVIDIA GeForce GTX 1070. The total learning took about 10 hours of computation time. Note that once the CNN has been trained, the network computation time reduces to a few seconds.

3.2.3 Network testing with the synthetic dataset

We evaluated the proper multiparametric mapping of the network through the 14 brain volumes of the synthetic dataset remaining for testing. In addition to visual evaluation, we carried out a quantitative analysis in the parametric maps domain due to the existence of the corresponding ground-truth. The comparison between the computed and the ground-truth T1, T2, and PD parametric maps was performed with the nomalized squared error (NSE) map computed as

$$\mathrm{NSE}(\mathbf{x}) = \frac{(\mathrm{MAP_c}(\mathbf{x}) - \mathrm{MAP_{GT}}(\mathbf{x}))^2}{\mathrm{MAP^2_{GT}}(\mathbf{x})} \times 100\%, \tag{3.1}$$

where MAP is one of the T1, T2, and PD maps, and the subscripts c and GT stand for *computed* and *ground-truth*, respectively. Similarly, the computed and ground-truth parametric maps are also compared with the scalar metrics described later in section 3.3.1.

3.2.4 Validation with actual MR acquisitions

Two different datasets of actual MR acquisitions are employed in this work, namely *Multicontrast Brain-MRI* and *Relaxometry Brain-MRI*, which have been previously described in Chapter 1. In the current chapter, the T1, T2, and PD maps of *Relaxometry Brain-MRI* will be referred to as silver standard, since they are affected by common artifacts as well as by physiological motion due to the length of their associated relaxometry sequences.

We preprocessed the actual MR brain volumes of these datasets in order to register all the image modalities to the same image space and to adapt them to the network input layer. All image modalities were affine-registered to the T2w-FLAIR image using FLIRT of FSL (Jenkinson et al., 2012). After registration, the size of each image modality is of 256 x 256 with 27 slices with voxel size of 0.94 x 1.25 x 5 mm as shown in Table A.2 of Appendix A. Note that this registration step is only necessary for training and validation purposes, because in production mode —once the network is fully trained— the only requirement is to have the input images with spatial alignment. All images were then skull-stripped. Subsequently, all images were cropped to 240 x 176 pixels which is the dimension of the network's input layer. We normalized the weighted images by dividing each of them by its average intensity without considering the background. This normalization was done in accordance with the preprocessing steps of the CNN training data. In addition, the relaxometry PD maps were normalized so that their 99th percentiles matched the maximum of the PD map from the synthetic training dataset. Finally, the 14 central slices of each actual MR brain volume were then selected to avoid slices with predominant background areas and/or very prone to artifacts.

We validated the performance of the proposed approach to compute parametric maps and to synthesize different weighted image modalities when actual T1w and T2w images are input to the network following the pipeline in Figure 3.1.b). Synthesis quality has been assessed both on the maps directly provided by the network output as well as on the synthesized weighted images. Quality parameters have been defined both at region of interest (ROI) level and at whole image level. Precise definitions for these parameters are provided in the next section.

3.3 Experimental work

3.3.1 Quantitative parameters for quality assessment

We drew nine circular ROIs in each subject of the *Multicontrast Brain-MRI* dataset co-localized across the different parametric maps and weighted images enumerated in Figure 3.1.b). From the nine ROIs, three were located in the CSF (approximately 3 mm of radius), three in the WM (approximately 3 mm of radius), and three in the GM (approximately 2 mm of radius). Let $\mathcal{X}_i^k(n)$ denote the set of voxels* belonging to ROI i, $1 \leq i \leq 3$ from tissue k, $1 \leq k \leq 3$ (say, 1 for CSF, 2 for GM, and 3 for WM) and subject n, $1 \leq n \leq 8$. As for the parametric maps provided by the network from *Multicontrast Brain-MRI*, we define the following two parameters

$$\mu^k = \frac{1}{\sum\limits_{n=1}^{8} \sum\limits_{i=1}^{3} |\mathcal{X}_i^k(n)|} \sum_{n=1}^{8} \sum_{i=1}^{3} \sum_{\mathbf{x} \in \mathcal{X}_i^k(n)} \mathrm{MAP}_{\mathrm{c}}^{n}(\mathbf{x}) \tag{3.2}$$

$$s^k = \sqrt{\frac{1}{\sum\limits_{n=1}^{8} \sum\limits_{i=1}^{3} |\mathcal{X}_i^k(n)|} \sum_{n=1}^{8} \sum_{i=1}^{3} \sum_{\mathbf{x} \in \mathcal{X}_i^k(n)} \left(\mathrm{MAP}_{\mathrm{c}}^{n}(\mathbf{x}) - \mu^k\right)^2} \tag{3.3}$$

with $\mathrm{MAP}_{\mathrm{c}}^{n}(\mathbf{x})$ a computed parametric map evaluated at point $\mathbf{x}$ and $|\cdot|$ denotes the cardinality of a set.

For the particular case of the subjects of *Relaxometry Brain-MRI* dataset, we drew 12 circular ROIs in each tissue co-localized across the different parametric maps and weighted images. For its parametric maps we define:

$$\mu_i^k(n)_{\mathrm{L}} = \frac{1}{|\mathcal{X}_i^k(n)|} \sum_{\mathbf{x} \in \mathcal{X}_i^k(n)} \mathrm{MAP}_{\mathrm{L}}^{n}(\mathbf{x}) \tag{3.4}$$

with $\mathrm{MAP}_{\mathrm{L}}^{n}(\mathbf{x})$ a parametric map of the n-th healthy subject evaluated at point $\mathbf{x}$ and L is a label that takes the values c for the MAP computed by the network and Silver for the silver standard relaxometry maps; k follows the same convention as in Eqs. (3.2) and (3.3), $1 \leq i \leq 12$, and $1 \leq n \leq 5$.

As for the weighted images of *Multicontrast Brain-MRI*, we define:

$$\mu_i^k(n) = \frac{1}{|\mathcal{X}_i^k(n)|} \sum_{\mathbf{x} \in \mathcal{X}_i^k(n)} m^n(\mathbf{x}) \tag{3.5}$$

$$s_i^k(n) = \sqrt{\frac{1}{|\mathcal{X}_i^k(n)|} \sum_{\mathbf{x} \in \mathcal{X}_i^k(n)} \left(m^n(\mathbf{x}) - \mu_i^k(n)\right)^2} \tag{3.6}$$

$$s(n) = \frac{1}{9} \sum_{i=1}^{3} \sum_{k=1}^{3} s_i^k(n) \tag{3.7}$$

---

*ROIs have been delineated in 2D, so the third component $\forall \mathbf{x} \in \mathcal{X}_i^k(n)$ coincides.

with $m^n(\mathbf{x})$ an image (either synthesized or acquired) of the n-th subject evaluated at point $\mathbf{x}$. Then, the following samples (per tissue k, $1 \leq k \leq 3$) are created:

1. Intensity values $\mu_i^k(n)$, $1 \leq i \leq 3$, $1 \leq n \leq 8$.
2. Contrast:
$$c_{ij}^k(n) = \frac{\mu_i^k(n) - \mu_j^k(n)}{\mu_i^k(n) + \mu_j^k(n)}, \tag{3.8}$$
$1 \leq i, j \leq 3$, $i \neq j$, $1 \leq n \leq 8$.
3. contrast-to-noise ratio (CNR):
$$\mathrm{CNR}_{ij}^k(n) = \frac{\mu_i^k(n) - \mu_j^k(n)}{s(n)}, \tag{3.9}$$
$1 \leq i, j \leq 3$, $i \neq j$, $1 \leq n \leq 8$.
4. signal-to-noise ratio (SNR):
$$\mathrm{SNR}_i^k(n) = \frac{\mu_i^k(n)}{s(n)}, \tag{3.10}$$
$1 \leq i \leq 3$, $1 \leq n \leq 8$.

In addition, in each subject of *Multicontrast Brain-MRI* we also drew a rectangular ROI measuring approximately 70.50 mm × 33.75 mm, which was chosen to encompass the occipital region of the brain. The number of pixels of this rectangular ROI was of 2025.

At a whole image level, we used four well-known metrics commonly used in medical image translation methods. These metrics are the mean squared error (MSE), the structural similarity index (SSIM), the peak signal-to-noise ratio (PSNR), and the correlation coefficient (CORR) defined as follows:

$$\overline{m^n} = \frac{1}{|\mathcal{X}|} \sum_{\mathbf{x} \in \mathcal{X}} m^n(\mathbf{x})$$

$$c_{m_1^n m_2^n} = \frac{1}{|\mathcal{X}|} \sum_{\mathbf{x} \in \mathcal{X}} \left(m_1^n(\mathbf{x}) - \overline{m_1^n}\right) \left(m_2^n(\mathbf{x}) - \overline{m_2^n}\right)$$

$$\mathrm{MSE}(n) = \frac{1}{|\mathcal{X}|} \sum_{\mathbf{x} \in \mathcal{X}} \left(m_{\mathrm{syn}}^n(\mathbf{x}) - m_{\mathrm{acq}}^n(\mathbf{x})\right)^2 \tag{3.11}$$

$$\mathrm{PSNR}(n) = 10 \log_{10} \left(\frac{\max\limits_{\mathbf{x} \in \mathcal{X}} \left(m_{\mathrm{syn}}^n(\mathbf{x})\right)^2}{\mathrm{MSE}(n)} \right) \tag{3.12}$$

$$\mathrm{CORR}(n) = \frac{c_{m_{\mathrm{syn}}^n m_{\mathrm{acq}}^n}}{\sqrt{c_{m_{\mathrm{syn}}^n m_{\mathrm{syn}}^n} c_{m_{\mathrm{acq}}^n m_{\mathrm{acq}}^n}}} \tag{3.13}$$

$$\mathrm{SSIM}(n) = \frac{(2\overline{m_{\mathrm{syn}}^n} \; \overline{m_{\mathrm{acq}}^n} + C_1)(2c_{m_{\mathrm{syn}}^n m_{\mathrm{acq}}^n} + C_2)}{\left(\left(\overline{m_{\mathrm{syn}}^n}\right)^2 + \left(\overline{m_{\mathrm{acq}}^n}\right)^2 + C_1\right) \left(c_{m_{\mathrm{syn}}^n m_{\mathrm{syn}}^n} + c_{m_{\mathrm{acq}}^n m_{\mathrm{acq}}^n} + C_2\right)} \tag{3.14}$$

with $m^n_{\text{syn}}(\mathbf{x})$ and $m^n_{\text{acq}}(\mathbf{x})$ the synthesized and acquired images, respectively, for the n-th subject of *Multicontrast Brain-MRI*, $1 \leq n \leq 8$, evaluated at point $\mathbf{x}$; voxels take on values within domain $\mathcal{X}$. Unless otherwise stated, this domain will consist in the brain area. These four metrics have also been used with parametric maps for performance assessment on synthetic data.

3.3.2 Experiments

Network verification with synthetic images as inputs has been accomplished by visual assessment as well as with the NSE map defined in Eq. (3.1). In addition, the parameters defined in Eqs. (3.11)—(3.14) have also been employed.

As for the network validation with real images, all the parameters defined in the previous section have been employed, and assessment has been carried out both directly on the network outputs (i.e., on the parametric maps) as well as on the synthesized weighted images. For the former, we have employed the silver standard maps from the five subjects included in the *Relaxometry Brain-MRI* dataset. For the latter, and as indicated in Figure 3.1.b), we analytically synthesized the same weighted images acquired in the *Multicontrast Brain-MRI* dataset with the same sequence parameters as those described in Table A.1. The equations used for each sequence are Eqs. (2.10) - (2.13) described in Subsection 2.1.4 for the synthesis of T1w, PDw/T2w, T2*w, and T2w-FLAIR, respectively.

In addition, we synthesized additional weighted images with the same sequences as in the *Multicontrast Brain-MRI* dataset, but varying the sequence parameters (i.e. TE, TR, TI). These sequences are SE (Eq. (2.11)) with TE in the range of 20 to 100 ms and TR of 120 and 4000 ms, and IR-SE (Eq. (2.13)) with three different combinations of TE, TR, and TI. We do not have the corresponding acquired weighted images as ground-truth due to scan time restrictions, but we pursue to investigate the versatility of our approach to synthesize any weighted image with coherent contrast.

We have also tested how the network deals with non skull-stripped images, a fact that is indicated to be an issue in Chartsias et al. (2017). To this end, non skull-stripped T1w and T2w images were input to the CNN. In this case, normalization was done by dividing each of them by the skull-stripped images average intensity in accordance with the synthetic dataset generation and network training. From the parametric maps with skull computed by the CNN, we then analytically synthesized the same weighted images included in the *Multicontrast Brain-MRI* dataset.

Finally, we propose a network refinement by performing additional training with a small number of real weighted images and their corresponding silver standard parametric maps of *Relaxometry Brain-MRI.* We have carried out a cross validation procedure; specifically, we tested with $5 - t$ subjects, where $2 \leq t \leq 4$, and the remaining t subjects have been divided into training and early-stopping monitoring datasets; cross validation stems from the fact that we have $\binom{5}{t}$ combinations of testing datasets for each t; each combination will be hereinafter referred to as a split. Note that the case $t = 4$ corresponds to a leave-one-out scheme. Within this scheme, we have carried out two experiments: (i) the CNN previously trained with the synthetic dataset is fined tuned with the parametric maps and (ii) the CNN is trained from scratch making use exclusively of the parametric maps of *Relaxometry Brain-MRI* (i.e., no synthetic data are shown to the CNN). Maps from experiment (i) will be referred to as $\mathrm{MAP}_{\text{c-(i)}}$ while maps from experiment (ii) will be denoted by $\mathrm{MAP}_{\text{c-(ii)}}$.

3.3.3 Statistical Analysis

The parameters defined in Eqs. (3.11)—(3.14) when applicable, are shown as averages (and sample standard deviation) along the 14 synthetic volumes used for testing or the 8 subjects used for system validation; these parameters are calculated within a 3D domain of the 14 central slices (where largest brain areas are found).

As for parameters defined in Eqs. (3.4), (3.5), (3.8)—(3.10) we have measured the Pearson correlation coefficient and the intra-class correlation coefficient (ICC), particularly ICC(2,1) (Koo and Li, 2016). As for the former, we ran a correlation test based on the Fisher transformation to test the hypothesis that the correlation coefficient is less than or equal to a predefined value; a p-value (p) < 0.05 was considered significant so as to reject the hypothesis. We have also analyzed Eqs. (3.8)—(3.10) using linear regression. Additionally, for the rectangular ROI drawn in the synthesized and acquired weighted images we have computed the Pearson correlation coefficient and performed an F-test for linear regression. Finally, we carry out a Bland-Altman plot analysis of a representative slice per subject where pixel values were normalized so that a value of “1.0” represented the signal strength of WM for each particular weighted image as in Cheng et al. (2020).

3.4 Results

3.4.1 Network testing with the synthetic dataset

Figure 3.2 shows a representative axial slice of the T1, T2, and PD maps computed from one of the test brain volumes of the synthetic dataset together with their corresponding NSE maps. Main differences between the computed and ground-truth maps appear in the boundary of the brain and in the tissue interfaces to a lower extent. Nevertheless, the NSE is predominantly below 1% on the three computed T1, T2, and PD maps. Further, the mean evaluation metrics obtained in the synthetic data testing of all 14 test brain volumes show good agreement between the computed and the ground-truth maps as can be seen in Table 3.2. The SSIM is always above 0.99 and the MSE below 1%.

3.4.2 Validation with actual MR acquisitions

Table 3.3 shows the parameters defined in Eqs. (3.2) and (3.3) for the T1, T2, and PD parametric maps obtained from all the ROIs within a tissue along all the subjects of the *Multicontrast Brain-MRI* dataset. The values obtained in this work are mostly within the range of the values previously reported in the literature for a

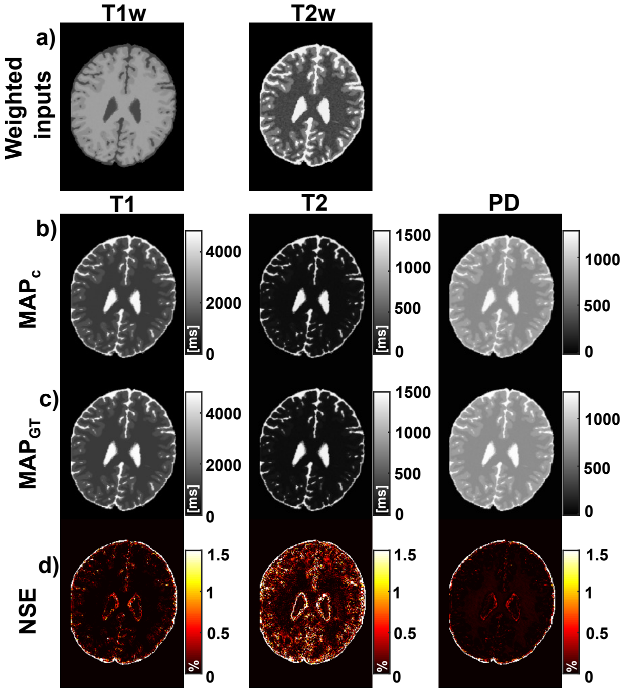


Figure 3.2: A representative axial slice of the T1, T2, and PD maps computed from a test brain volume of the synthetic dataset. a) T1w and T2w images input to the network. b-d) Computed, ground-truth, and NSE maps from the same slice for the T1, T2, and PD parameter maps, respectively. The T1 and T2 values are given in miliseconds (ms). Main differences between the computed and the ground-truth maps appear in the boundary of the brain, although the NSE is predominantly below 1% on the three computed T1, T2, and PD maps.

Table 3.2: Metrics (mean ± SD) used to evaluate the performance of the CNN to compute each set of T1, T2, and PD maps from each pair of T1w and T2w images of the test brain volumes of the synthetic dataset. These metrics are the mean squared error (MSE), structural similarity error index (SSIM), peak signal-to-noise ratio (PSNR), and correlation coefficient (CORR). The metrics were calculated between the computed parametric maps and the ground-truth T1, T2, and PD maps. Note that for the calculation of the metrics the background voxels were not considered.

	T1	**T2**	**PD**
MSE	0.0072 (0.0044)	0.0013 (0.0010)	0.0004 (0.0002)
SSIM	0.9932 (0.0016)	0.9933 (0.0044)	0.9912 (0.0033)
PSNR	36.1274 (2.2800)	33.8001 (2.5282)	37.2614 (1.5868)
CORR	0.9983 (0.0007)	0.9975 (0.0007)	0.9990 (0.0004)

3T scanner. As for the particular case of the PD maps, the GM/WM ratio is close to the ratio reported in the literature (1.22 vs. 1.10, respectively). In addition, Figure 3.3 shows a representative axial slice of the T1, T2, and PD maps computed from a subject of the *Relaxometry Brain-MRI* dataset and their corresponding silver standard relaxometry maps. The computed parametric maps are visually realistic and capture most of the structural information without computational errors. Note that no outliers appear in the CSF of the T1 map. Computed parametric maps present less blurring than the silver standard relaxometry maps, this could be explained by the fact that neither inhomogeneity nor motion corrections were applied to the estimation of the latter. The correlation diagrams include the values of parameter $\mu_i^k(n)_L$ (Eq. 3.4) for the five subjects of the *Relaxometry Brain-MRI* dataset. There is high correlation between the computed and the silver standard relaxometry maps, namely 0.9616, 0.9703, and 0.7707 for the T1, T2, and PD, respectively; the first two values are statistically higher than 0.90 ($p < 0.01$). Similarly, respective ICC values are 0.9454, 0.9445, and 0.6489.

Table 3.3: Values defined in Eqs. 3.2 and 3.3 (the latter, within braces) for each tissue in each of the computed parametric maps (i.e. T1, T2, and PD maps). Comparison with the values previously reported in the literature for a 3T scanner.

	T1 (s)		**T2 (s)**		**PD**	
	This work	**Literature**	**This work**	**Literature**	**This work**	**Literature**
WM	0.9741 (0.0585)	0.7370 - 1.1000 [Zhu and Penn (2005), Deoni (2007)]	0.0890 (0.0061)	0.0560 - 0.0840 [Gelman et al. (1999), Wansapura et al. (1999)]	0.7222 (0.0142)	0.6330 [Hagiwara et al. (2019a)]
GM	1.4474 (0.1361)	1.3310 - 1.8200 [Wansapura et al. (1999), Stanisz et al. (2005)]	0.1257 (0.0160)	0.0710 - 0.1320 [Gelman et al. (1999), Wansapura et al. (1999)]	0.7988 (0.0135)	0.7720 [Hagiwara et al. (2019a)]
CSF	4.6785 (0.1060)	3.7000 - 6.8730 [Liberman et al. (2014), Clare and Jezzard (2001)]	1.3705 (0.0335)	0.5000 - 1.8700 [Deoni et al. (2004), Piechnik et al. (2009)]	1.2601 (0.0380)	-

Figure 3.4 shows a representative axial slice of weighted images synthesized from one set of the T1, T2, and PD maps computed by the CNN and their corresponding acquired images for a subject of the *Multicontrast Brain-MRI* dataset. Overall, the synthesized and acquired weighted images are visually similar regarding both structural information and contrasts between tissues. The image modalities used to train the network present higher similarity than the others, being the T1w the most similar and the T2w-FLAIR the least similar but yet with visual resemblance. The boundary of the CSF on the cortical area is hyperintense on the synthesized T2w-FLAIR, which is presumably caused by partial volume effects.

Figure 3.5 shows a scatter plot between the synthesized and the acquired weighted images (we show the pairs of values $\mu_i^k(n)$ defined in Eq. (3.5) for the acquired and the synthesized images) for the eight subjects of *Multicontrast Brain-MRI*. There is high correlation between the pairs of weighted images, namely 0.9979, 0.9952, 0.9912, 0.9820, and 0.9602 for the T1w, T2w, PDw, T2*w, and T2w-FLAIR, respectively; all of these values are statistically higher than 0.90 ($p < 0.001$). Similarly, respective ICC values are 0.9918, 0.9438, 0.8919, 0.7424, and 0.9367.

Figure 3.6 shows similar scatter plots of the contrast, CNR, and SNR samples between the synthesized and acquired weighted images (as defined in Eqs. (3.8)—

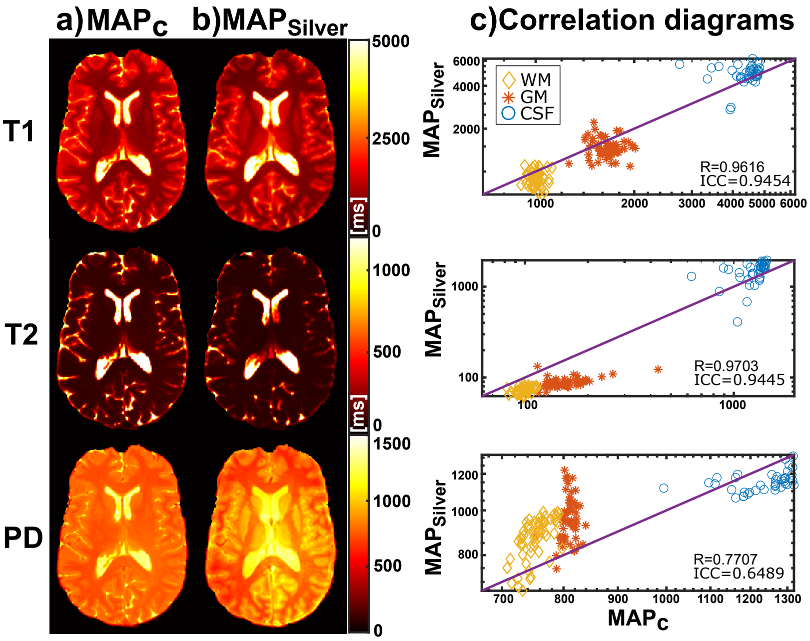


Figure 3.3: A representative axial slice of T1, T2, and PD maps computed from a subject of the *Relaxometry Brain-MRI* dataset. a) Computed T1, T2, and PD parametric maps. b) Their corresponding silver standard relaxometry maps. c) Correlation of parameter $\mu_i^k(n)_L$ (Eq. 3.4) between the computed and the silver standard relaxometry maps for the five healthy subjects. T1 and T2 values are given in miliseconds (ms). The markers indicate the mean values of WM (yellow diamonds), GM (red stars), and CSF (blue circles). Diagonal lines represent the identity.

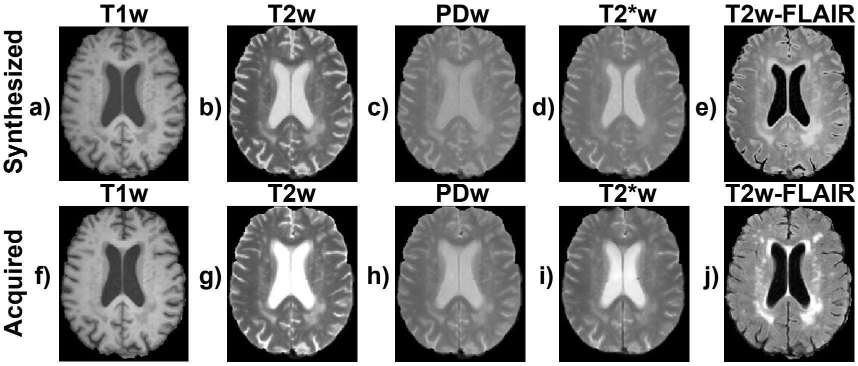


Figure 3.4: A representative axial slice of the weighted images synthesized from one set of the T1, T2, and PD maps computed by the CNN and their corresponding acquired images. a-e) The synthesized T1w, T2w, PDw, T2*w, and T2w-FLAIR images. f-j) Their corresponding weighted acquired images.

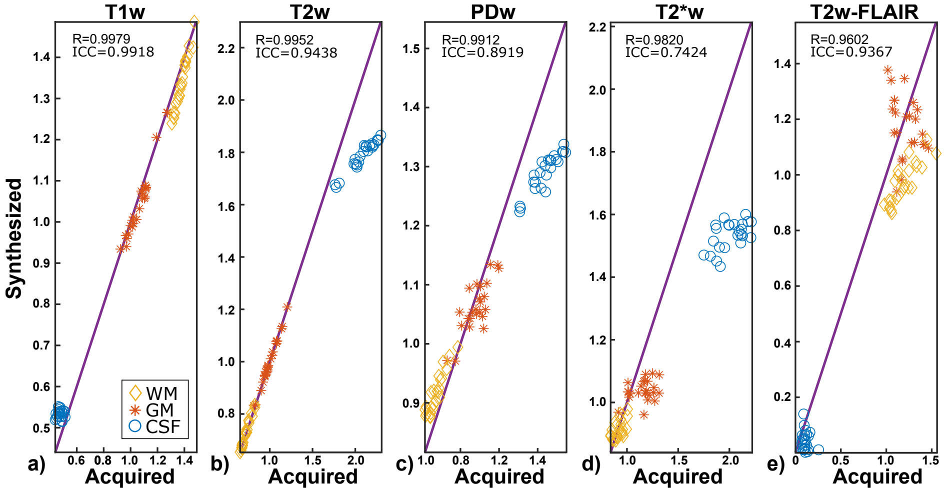


Figure 3.5: Correlation of parameter $\mu_i^k(n)$ (Eq. 3.5) between the synthesized and the acquired weighted images. a) T1w, b) T2w, c) PDw, d) T2*w, and e) T2w-FLAIR. The markers indicate the mean values of WM (yellow diamonds), GM (red stars), and CSF (blue circles). The diagonal lines represent the identity.

(3.10)). There is high correlation with values between 0.9907 and 0.9241 for the contrast, 0.9807 and 0.8739 for the CNR, and 0.9845 and 0.9082 for the SNR for all the weighted images, as shown in Table 3.4; most of these values are statistically larger than 0.90 ($p < 0.05$), except in the case of the T2*w and T2w-FLAIR for the CNR and SNR, and the PDw only for the SNR, which are statistically greater than 0.84. The ICC values for the three same samples are between 0.9730 and 0.9421 for the T1w, between 0.9579 and 0.8954 for the T2w, and between 0.9120 and 0.8607 for the T2w-FLAIR. In contrast, the ICC values are lower for the PDw and T2*w, as shown in Table 3.4.

Table 3.4: Correlation coefficient (R) and intraclass correlation coefficient (ICC) of the contrast, the contrast-to-noise ratio (CNR), and the signal-to-noise ratio (SNR) between the synthesized and the acquired weighted images (see Figure 3.6). The bold correlation values indicate that they are statistically significant superior to a correlation value of 0.9. The * indicates $p < 0.05$ and ** $p < 0.001$.

		T1w	**T2w**	**PDw**	**T2*w**	**T2w-FLAIR**
Contrast	R	**0.9907****	**0.9855****	**0.9591****	**0.9241***	**0.9689****
	ICC	0.9421	0.9155	0.6521	0.6530	0.8961
CNR	R	**0.9807****	**0.9658****	**0.9453****	0.9193	0.8739
	ICC	0.9730	0.8954	0.7425	0.9187	0.8607
SNR	R	**0.9845****	**0.9734****	0.9280	0.9082	0.9145
	ICC	0.9712	0.9579	0.8300	0.4634	0.9120

Linear regression showed that the SNR of the synthesized weighted images is generally better with an improvement that reaches 47.74% (CI: [41.93%; 53.54%]). CNR is fairly similar for the T1w (-5.31%, CI: [-6.33%; -4.29%]), the T2*w (-0.12%,

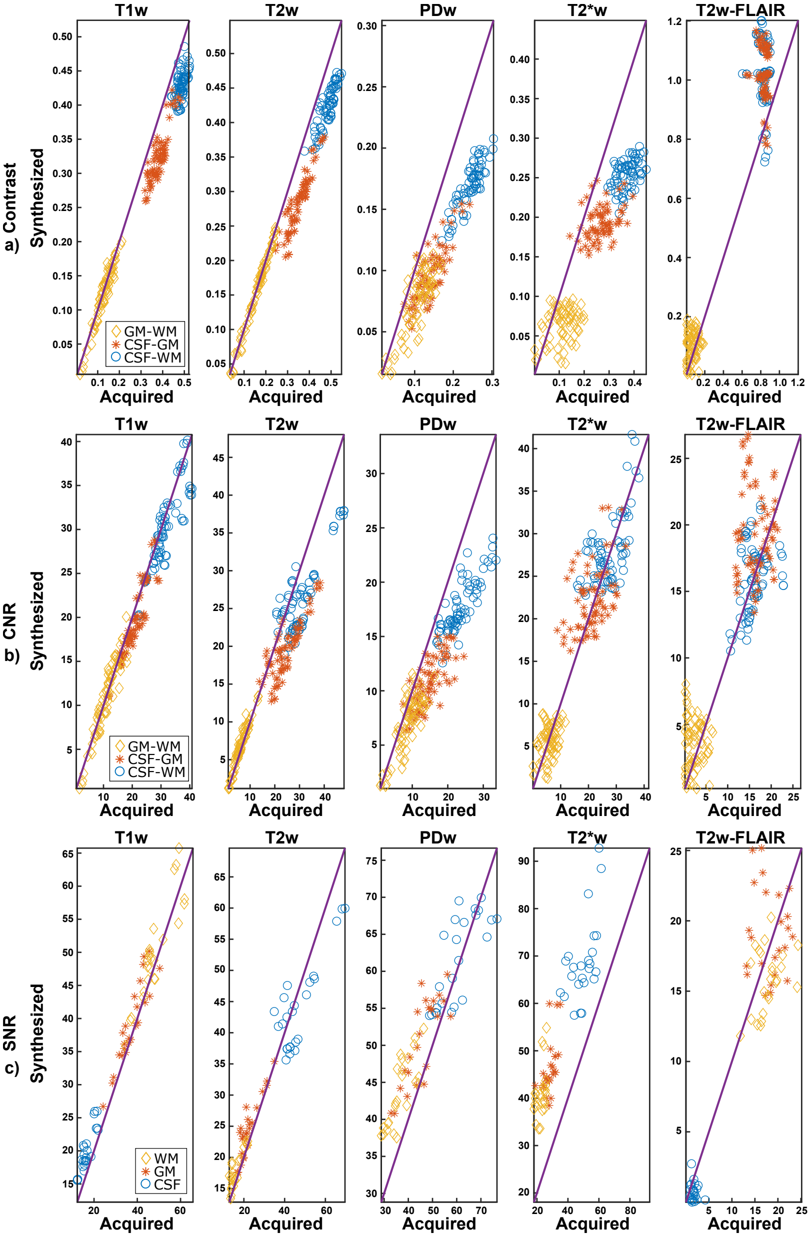


Figure 3.6: Correlation of the contrast, the CNR, and SNR between the synthesized and the acquired weighted images. a) Contrast, b) CNR, and c) SNR of the T1w, T2w, PDw, T2*w, and T2w-FLAIR images. For the Contrast (a) and the CNR (b) the markers indicate the contrast/CNR values between each combination of the GM ROIs with the WM ROIs (yellow diamonds), each combination of CSF ROIs with the GM ROIs (red stars), and each combination of CSF ROIs with the WM ROIs (blue circles). For the SNR (c) the markers indicate the mean SNR values of WM (yellow diamonds), GM (red stars), and CSF (blue circles). The diagonal lines represent the identity.

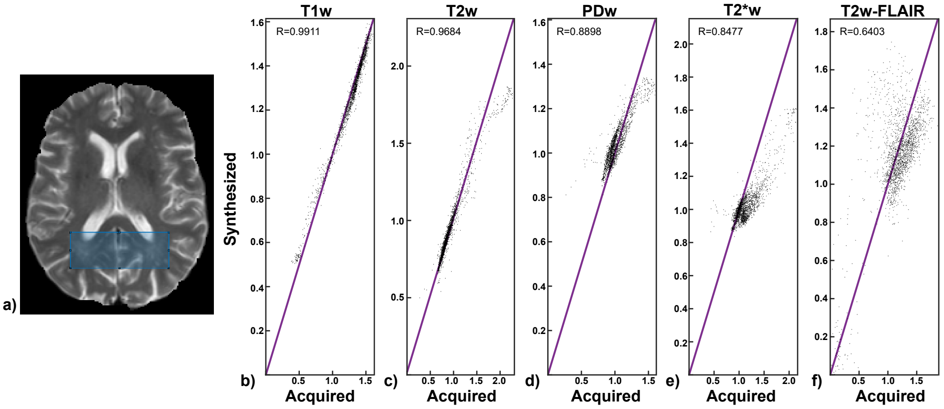


Figure 3.7: Correlation between pixel values from the rectangular ROI of the synthesized and the acquired weighted images for a representative subject of the *Multicontrast Brain-MRI* dataset. a) Rectangular region of interest chosen for linear regression, b) T1w, c) T2w, d) PDw, e) T2*w, and f) T2w-FLAIR. The markers indicate the value of each pixel in the rectangular ROI (approximate 2025 values). The diagonal lines represent the identity. p << 0.0001 for all image modalities for the correlation test carried out (see correlation values tested in main text).

CI: [-2.75%; 2.52%]), and the T2w-FLAIR (5.36%, CI: [1.67%; 9.05%]), although it is slightly worse for the T2w (-17.46%, CI: [-18.94%; -15.99%]) and the PDw (-26.45%, CI: [-27.79%; -25.10%]). Finally, contrast only improves in the T2w-FLAIR (22.77%, CI: [20.57%; 24.98%]). See details in Table 3.5.

Table 3.5: Percentage of variation of the linear regression coefficient [95% confidence interval (CI)] in comparison to the identity (i.e., linear regression coefficient of one) for the contrast, the CNR, and SNR samples. For the linear regression computation, the x-axis is considered as the samples values of the acquired image and the y-axis the samples values of the synthesized image as shown in Figure 3.6. Positive values indicate an improvement of the corresponding samples.

	T1w	**T2w**	**PDw**	**T2*w**	**T2w-FLAIR**
Contrast	-12.69% [-13.36; -12.01]%	-14.11% [-14.88; -13.33]%	-31.22% [-32.23; -30.21]%	-32.89% [-34.37; -31.34]%	22.77% [20.57; 24.98]%
CNR	-5.31% [-6.33; -4.29]%	-17.46% [-18.94; -15.99]%	-26.45% [-27.79; -25.10]%	-0.12% [-2.75; 2.52]%	5.36% [1.67; 9.05]%
SNR	4.76% [2.89; 6.62]%	-3.10% [-5.92; -0.28]%	6.86% [4.22; 9.51]%	47.74% [41.93; 53.54]%	-5.76% [-11.08; -0.44]%

Figure 3.7 shows a scatter plot between the pixel values of the rectangular ROIs drawn on the synthesized and the acquired weighted images of a representative subject of *Multicontrast Brain-MRI*. There is high correlation between the pairs of weighted images, namely 0.9911, 0.9684, 0.8898, 0.8477, and 0.6403 for the T1w, T2w, PDw, T2*w, and T2w-FLAIR, respectively; all of these values are statistically significant ($p << 0.0001$) in the F-test for linear regression.

Additionally, the high values of the mean SSIM, PSNR, and CORR and low values of the MSE obtained in the subjects of *Multicontrast Brain-MRI* show good agreement between synthesized and acquired weighted images as shown in Table 3.6. Specifically, the SSIM achieves values above 0.96 for the T1w and the T2w, and of 0.91, 0.78, and 0.56 for the PDw, T2*w, and T2w-FLAIR, respectively. The MSE is below 1% for the T1w, T2w, and PDw, and below 9% for the T2*w and the T2w-FLAIR. Similarly to Figure 3.4, the image modalities used to train the network show higher SSIM, PSNR, and CORR and lower MSE than the others.

Table 3.6: Metrics (mean ± SD) used to evaluate the capability to synthesize weighted images from a set of T1, T2, and PD maps computed by the CNN. The metrics were calculated between both the synthesized and the acquired weighted images. Note that for the calculation of the metrics the background voxels were not considered.

	T1w	**T2w**	**PDw**	**T2*w**	**T2w-FLAIR**
MSE	0.0058 (0.0009)	0.0095 (0.0020)	0.0061 (0.0010)	0.0392 (0.0058)	0.0815 (0.0081)
SSIM	0.9651 (0.0051)	0.9620 (0.0039)	0.9194 (0.0078)	0.7823 (0.0222)	0.5693 (0.0190)
PSNR	30.6338 (1.5330)	26.2621 (0.7607)	25.3160 (0.7972)	18.9098 (0.5905)	19.6598 (1.9280)
CORR	0.9910 (0.0015)	0.9858 (0.0023)	0.9886 (0.0017)	0.9438 (0.0076)	0.8726 (0.0093)

Figure 3.8 represents Bland-Altman plots including data from a representative slice of all subjects of *Multicontrast Brain-MRI.* It compares synthesized and acquired pixel values for T1w, T2w, PDw, T2*w, and T2w-FLAIR. The absolute mean difference for the each image modality is 0.0041, 0.0021, 0.0089, 0.0794 and 0.0249, respectively.

Figure 3.9 displays a representative axial slice of additional weighted images synthesized with the same sequences as in *Multicontrast Brain-MRI*, but varying the sequence parameters. This proves the versatility of the proposed approach to synthesize any weighted images. The images obtained are realistic and with coherent contrasts.

Finally, Figure 3.10 shows a representative axial slice of the non skull-stripped weighted images synthesized from one set of the T1, T2, and PD maps computed by the CNN and their corresponding non skull-stripped acquired images for a subject of *Multicontrast Brain-MRI.* Similarly to Figure 3.4, both images are visually apparent regarding both structural information and contrast between tissues. Nevertheless, the inhomogeneities in the skull interfaces might cause a mismatch between the synthesized and the acquired images.

3.4.3 Fine tuning: refining the network with actual parametric maps

Figure 3.11 shows a representative slice of both $\text{MAP}_{\text{c-(ii)}}$ and $\text{MAP}_{\text{c-(i)}}$ (columns a) and b), respectively) with MAP as T1, T2, and PD (first, second, and third

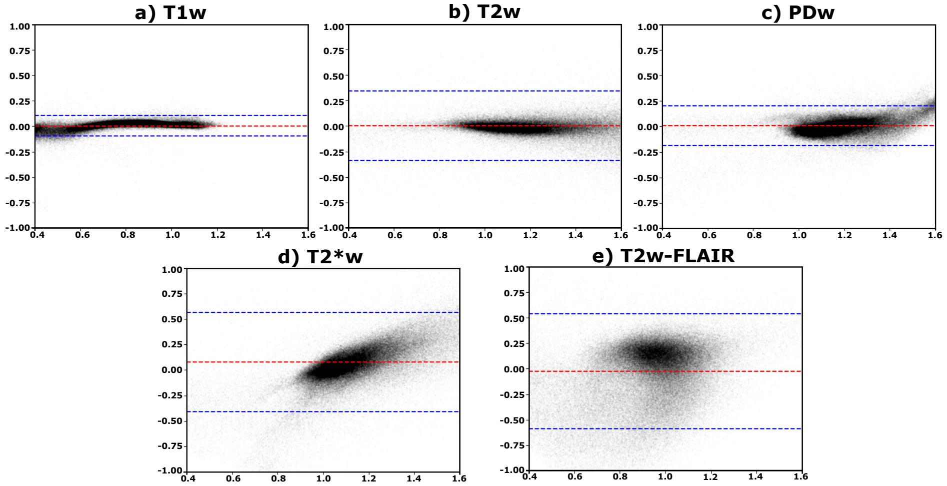


Figure 3.8: Bland-Altman plots used to compare synthesized and acquired weighted images as in Cheng et al. (2020). Each plot combines results from a representative axial slice of all subjects of the *Multicontrast Brain-MRI* dataset. The corresponding image modalities are: a-e) T1w, T2w, PDw, T2*w, and T2w-FLAIR, respectively. Red dashed lines represents the bias and blue dashed lines the 95% confidence interval.

rows, respectively). For $\text{MAP}_{\text{c-(i)}}$ the figure also shows the correlation diagrams that include the values of the ROIs for all subjects of the *Relaxometry Brain-MRI* dataset tested with a leave-one-out scheme (i.e., $t = 4$). It can be seen that the fine tuning procedure improves the accuracy of the computed parametric maps in terms of ICC (compare the values shown in Fig. 3.3), whereas without the previous synthetic training the results worsen noticeably and the maps blur. Furthermore, Table 3.7 shows the mean correlation coefficient and ICC of parameter $\mu_i^k(n)_L$ (Eq. 3.4) for the different configurations of t. Both correlation and ICC have been computed for each test subject of each split, and then, mean values were computed along all splits and subjects. Results show that, as expected, both parameters increase with the number of training subjects.

3.5 Discussion

In this chapter, we have presented a novel joint Synthetic MRI approach for the computation of the T1, T2, and PD parametric maps and the synthesis of different weighted images from only a pair of input weighted images. The pair of input images are a T1w and a T2w acquired with clinical routine sequences. The parametric maps are obtained with a DL method based on a CNN. This CNN is trained by means of a new training strategy with a synthetic dataset; hence, we overcome the lack of a public and sufficiently large database of conventional images that should be accompanied by their corresponding parametric maps.

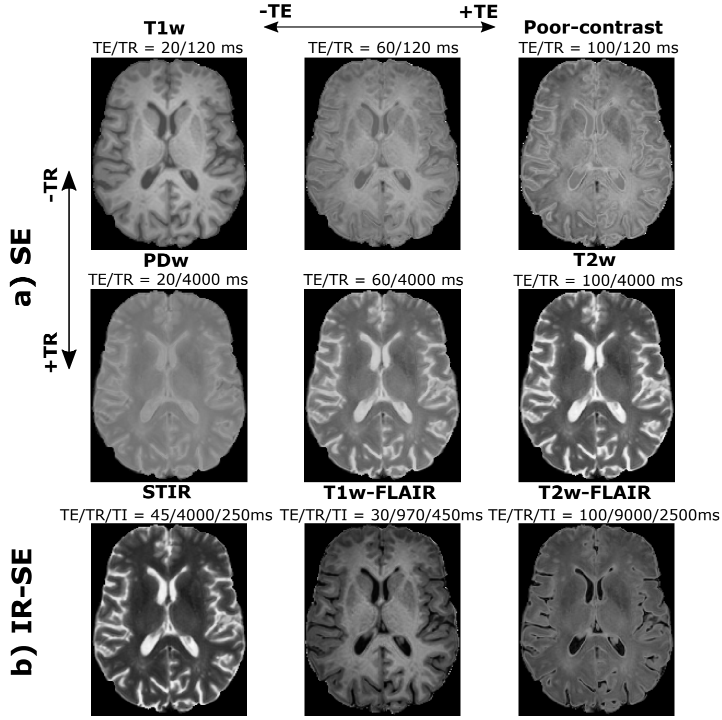


Figure 3.9: A representative axial slice of other weighted images synthesized varying the sequence parameters from a set of the T1, T2, and PD maps computed by the CNN. a) Weighted images synthesized for a spin echo (SE) sequence with different TE and TR corresponding to T1w, T2w and PDw image modalities. b) Weighted images synthesized for an IR-SE sequence with different TE, TR, and TI corresponding to short-TI inversion recovery (STIR), T1-weighted fluid attenuated inversion recovery (T1w-FLAIR), and T2w-FLAIR image modalities. Note that the unlabeled images correspond to sequence parameter combinations which lead to weighted images with undefined contrast.

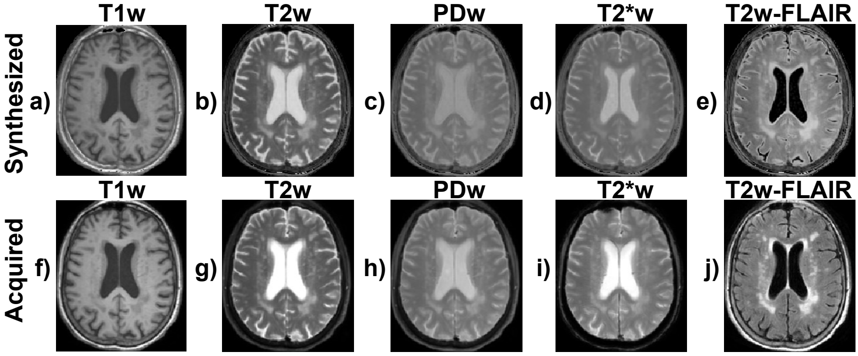


Figure 3.10: A representative axial slice of the non skull-stripped weighted images synthesized from one set of the T1, T2, and PD maps computed by the CNN and their corresponding non skull-stripped acquired images. a-e) The synthesized T1w, T2w, PDw, T2*w, and T2w-FLAIR images. f-j) Their corresponding weighted acquired images.

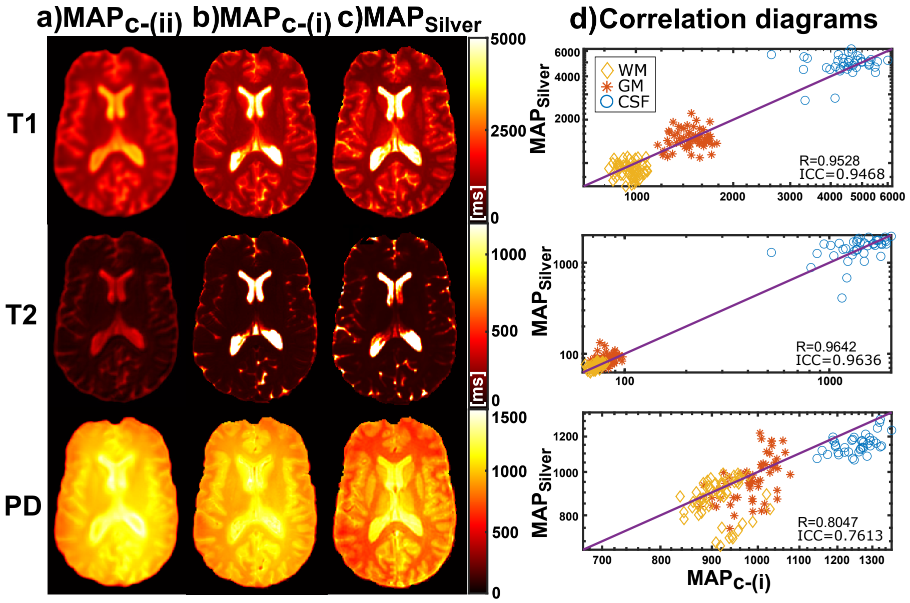


Figure 3.11: A representative axial slice of T1, T2, and PD maps computed from a subject of the *Relaxometry Brain-MRI* dataset in a leave-one-out scheme. a) T1, T2, and PD parametric maps computed by the network trained from scratch with actual parametric maps. b) Corresponding maps computed by the fine tuned network with previous synthetic training. c) Their corresponding silver standard relaxometry maps. d) Correlation of parameter $\mu_i^k(n)_L$ (Eq. 3.4) between $\mathrm{MAP}_{\text{c-(i)}}$ and $\mathrm{MAP}_{\text{Silver}}$ for the five healthy subjects. Diagonal lines represent the identity.

Table 3.7: Mean correlation coefficient (R) and intraclass correlation coefficient (ICC) of parameter $\mu_i^k(n)_L$ (Eq. 3.4) between the computed and the silver standard relaxometry maps for three values of the number of training subjects t in the cross validation of the fine tuning. Values reported have been computed for each test subjects of each split, and then, mean values were computed along all splits and subjects.

		T1	**T2**	**PD**
R	t=4*	0.9784	0.9682	0.8912
	t=3†	0.9733	0.9648	0.8857
	t=2†	0.9722	0.9645	0.8752
ICC	t=4*	0.9517	0.9607	0.7935
	t=3†	0.9324	0.9496	0.7930
	t=2†	0.9275	0.9475	0.7536

* five splits, † 10 splits.

Our synthetic training dataset departs from 120 instances of BrainWeb maps, in which we add intensity variability, by means of random noise, as well as spatial variability, by registering these maps to different anatomies from the PPMI database. We show the feasibility of this solution by computing accurate and realistic parametric maps from both synthetic and actual MR brain acquisitions; the computed maps are then used to synthesize different weighted images, so our end-to-end Synthetic MRI solution is not limited to a number of predefined weighted images that have entered the learning process, but is capable of generalizing to any image modality that can be synthesized out of the parametric maps. Hence, our solution fulfills the three conditions needed to become a Synthetic MRI method. To the best of our knowledge this is the first Synthetic MRI method that is based on conventional routine sequences and can be trained on the basis of synthetic data.

We have shown that synthesized weighted images from five clinical routine sequences achieve high similarity metrics, with SSIM usually above 0.90 and low error with MSE always below 9%. The correlation analysis shown in the scatter plots of Figure 3.5 provide values above 0.95 for all modalities. Similarly, for the scatter plots of contrast, CNR, and SNR (Figure 3.6), both correlation and ICC also obtain high values, as shown in Table 3.4. Note that the agreement when the ICC values are above 0.75 is considered good while when the values are above 0.90 is considered excellent (Koo and Li, 2016); our results indicate that we lie in these ranges for at least one parameter for each synthesized modality. In addition, spatial resolutions of training and test images do not need to exactly match. Our testing images have resolution of 0.94 x 1.25 x 5 mm while the PPMI datase resolution is of 1 x 1 x 1.2 mm; despite the in-plane resolution does not differ much, slice thicknesses are clearly different and no partial volumes effects in the through-plane direction are obvious in our solution.

The Synthetic MRI approach proposed may have important implications in neuroimaging due to the utility of the parametric maps for tissue characterization and the possibility of synthesizing any weighted image. Specifically, the obtained T1, T2, and PD values of the three tissues (WM, GM, and CSF) present a good correspondence with the values reported in the literature, as shown in Table 3.3 and with the silver standard relaxometry parametric maps with correlation values above 0.95 for the T1 and T2 maps. The output quality increases noticeably when the network is fine tuned with a small number of silver standard maps. This provides a way to obtain parametric maps with increased accuracy, at the cost of employing a (small) number of silver standard maps for additional training. Note, however, that training with synthetic data is a key step, since training from scratch with this small amount of silver standard maps by no means suffices.

Moreover, the proposed approach avoids the need of lengthy relaxometry sequences; the total scan time of the full-brain acquisition described in this chapter (T1w and T2w acquisitions) is less than 8 min versus the 18 min scan time of an inversion recovery golden standard acquisition only for T1 mapping (Ramos-Llorden et al., 2016), and the 17 min scan time of the DESPOT algorithm for T1 and T2 mapping (Deoni et al., 2005). The computed parametric maps are therefore less prone to motion artifacts. Interestingly, the proposed approach is not based on specific and complex sequences as MR Fingerprinting (Ma et al., 2013), IR-TrueFISP (Gulani et al., 2004), QRAPMASTER (Warntjes et al., 2008), and MPME (Cheng et al., 2020) or private protocols as the SyMRI IMAGE software (SyntheticMR). Also, the feasibility of synthesizing weighted images and/or retrospectively optimizing sequence parameters can further reduce scan time. Thus, a radiologist could have the parametric maps together with various conventional weighted images based on the same widespread short scan protocol.

In addition, we provide the possibility of creating databases of perfectly registered weighted images accompanied with their corresponding parametric maps; these databases can be used to train machine learning algorithms for different purposes, perform data augmentation or improve the performance of registration or segmentation algorithms. The field of radiomics also seems a natural target for our methodology.

We should stress that our method gives rise to different modalities, some of them unseen by the network throughout the training process, with comparable quality with recent medical image translation works; however, to the best of our knowledge, these works are limited to the specific modalities that enter the training and validation stages. Specifically, SSIM in our synthesized T1w, T2w, and PDw images is slightly higher than the values reported by Chartsias et al. (2017) and Sohail et al. (2019) for some of these image modalities, albeit the T2w-FLAIR and sometimes the T2*w achieve lower quality. A more thorough comparison is not feasible since our actual acquired validation dataset is not large enough to train the state-of-the-art medical image translation methods. Also, note that in our approach no data from real acquisitions are used in the training stage thanks to the synthetic training, and only when fine tuning the network a very small database of actual maps is used; this is our main advantage with respect to the state-of-the-art.

The non skull-stripped synthesized images, although visually realistic, achieve lower quality than their corresponding skull-stripped counterparts; however, this seems to be the case as well in Chartsias et al. (2017). The loss of quality is clearer in the neighbouring parts of the skull, and other tissues such as the eyes. We obtain a SSIM of 0.80 in both the T2w and the PDw images while in Chartsias et al. (2017) the SSIM in the synthesis of the T2w from the PDw is of 0.86. Nevertheless, it is important to note that, as opposed to Chartsias et al. (2017), in our work the skull has not entered the training process.

As for the comparison with Synthetic MRI methods, in our work the synthesized weighted images show higher visual resemblance to the acquired images than in the other methods (Cheng et al., 2020; Gulani et al., 2004). In addition, correlation coefficients calculated within the rectangular ROIs described in section 3.3.1 are higher with our approach (Gulani et al., 2004). However, except for the T2w-FLAIR modality, our synthesized weighted images present lower contrast values than the acquired weighted images as compared with Blystad et al. (2012), although we achieve similar or higher CNR and SNR figures. The Bland-Altman plots show better agreement than Cheng et al. (2020) in T1w and T2w, similar agreement in PDw, and only a slightly lower agreement in T2w-FLAIR images. The T2*w modality is not synthetized by them. The loss of quality in T2w-FLAIR images is a common issue in Synthetic MRI (Hagiwara et al., 2017) where the boundary of the CSF on the cortical area tends to be hyperintense presumably due to partial volume effects.

This work has several limitations. The method was evaluated in synthetic data, eight subjects —suspected of early Alzheimer disease—, and five healthy volunteers, so further validation in a larger cohort of both healthy volunteers and patients with other pathologies is still needed. Additionally, silver standard relaxometry maps were estimated without applying neither inhomogeneity nor motion corrections. Moreover, B0 and B1 inhomogeneties have not been taken into account in the synthetic dataset generation either. Thus, including these inhomogeneties in the synthesis could be of interest. In addition, the equations used to synthesize the weighted images did not consider all the effects that occur in practice. For example, the T2*w image is synthesized from the T2 map instead of the T2* map which could be the cause of the worse metrics compared to the metrics of the T2w image. Also, the T2w-FLAIR presents worse metrics than the other modalities, but the images obtained are comparable with those of the literature (Blystad et al., 2012). To address this, methods focused on improving T2w-FLAIR images have been reported (Hagiwara et al., 2019b). Additionally, the parametric maps of the synthetic dataset were generated with values corresponding to 3T scanners; hence our results do not directly carry over to other field strengths. The extension to high field scanners will presumably require to modify the synthetic training dataset and further postprocessing corrections because B1 and B0 field inhomogeneities are specially problematic at high fields.

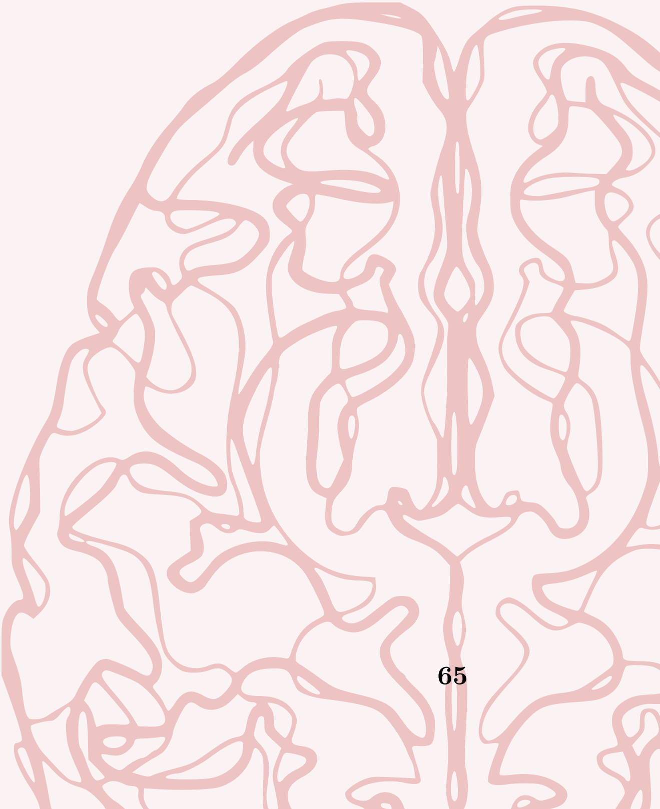
Future work includes improvement and further tuning in the implemented CNN. In addition, the use of GAN architectures may be studied due to the recent works that achieve impressive results in medical image translation capturing high-frequency texture information (Armanious et al., 2020; Dar et al., 2019). On the other hand, the selection of the optimal input training sequences and/or sequence parameters could improve the computation of the parametric maps and, subsequently, the synthesis of the weighted images. A more realistic synthesis of weighted images with a detailed Bloch simulation (Cao et al., 2014; Stöcker et al., 2010) and other maps (e.g. T2*, B0, and B1 maps) could also enhance the synthesis quality of any MRI modality. Further, the simulation of motion artifacts in the weighted images of the synthetic dataset could enhance the network robustness against such artifacts. In addition, we could extend the proposed approach to other slice orientations, such as sagittal or coronal, and/or other tissues, such as the heart or the liver, as long as synthetic anatomical volumes can be computed. To this end, for example, the extended Cardiac-Torso (XCAT) phantoms (Segars et al., 2010) could be employed. We also plan to address the oncology field, where parametric mapping may be a challenging task.

3.6 Conclusions

In conclusion, in this chapter we proposed a novel joint Synthetic MRI approach for the computation of the T1, T2, and PD parametric maps and the synthesis of different weighted images which only needs two conventional weighted images as inputs (full-brain acquisition in less than 8 min of scan time). Based on a CNN, we are able to provide realistic parametric maps and weighted images when training the CNN with a synthetic dataset. The results in both synthetic data and actual MR acquisitions experiments demonstrate its feasibility for quantitative MRI in clinically viable times as well as its applicability to the synthesis of additional MR weighted image modalities.

Part III

Synthetic MRI contributes to improvement of survival prediction



4

Synthetic MRI improves radiomics-based glioblastoma survival prediction

Contents

The work in this chapter has been published in:

- **Moya-Sáez, E.**, Navarro-González, R., Cepeda, S., Pérez-Núñez, A., de Luis-Garcia, R., Aja-Fernández, S., and Alberola-López, C. Synthetic MRI improves radiomics-based glioblastoma survival prediction. *NMR in Biomedicine.* 2022; 35(9), e4754.

- **Moya-Sáez, E.**, de Luis-Garcia, R., and Alberola-López, C. A self-supervised deep learning approach to synthesize weighted images and T1, T2, and PD parametric maps based on MR physics priors. *ISMRM & SMRT Annual Meeting & Exhibition, An Online Experience*, Virtual, May 2021; 2169.

- Navarro-González, R., **Moya-Sáez, E.**, de Luis-Garcia, R., Aja-Fernández, S., and Alberola-López, C. Synthetic MRI aids in glioblastoma survival prediction. *Joint ISMRM-ESMRMB & SMRT 31st Annual Meeting*, London, England, UK, May 2022; 3928.

4.1 Purpose

In this chapter we propose the application of Synthetic MRI to improve a radiomics approach for survival prediction in glioblastoma. Our purpose is to show that a radiomic system that incorporates an input channel fed by a synthesized image (a) behaves similarly to this system when it is fed with an acquired image and (b) undoubtedly outperforms a radiomics system that does not have this channel. Two weighted images are considered for the synthesis, namely, T2w and T2w-FLAIR.

We synthesize these images by means of a self-supervised extension of our DL Synthetic MRI approach presented in Chapter 3. The self-supervised training allows us to train the DL method with actually acquired weighted images instead of parametric maps, which are more difficult to obtain in practice. Hence, we validate an MR protocol shortening procedure by means of a glioblastoma survival prediction Radiomics-based application.

4.2 Methods

A total of 199 glioblastoma patients included in four different datasets were used in this work. Three of them (*BraTS2020*, *TCIA*, *Dataset22*), described in Appendix C.1, were only used for training the radiomics system. The other dataset (*Multicontrast Glioblastoma*), which has been previously described in Chapter 1, was used for testing the radiomics system in concordance with the synthesis method. Figure 4.1 shows the pipeline of the proposed approach for training and testing as well as and the three experiments we have carried out.

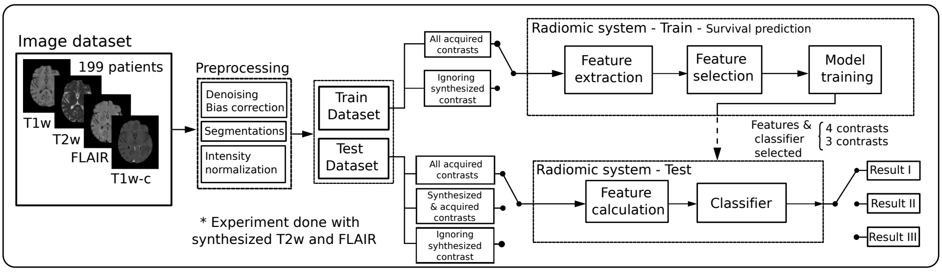


Figure 4.1: Workflow of the proposed approach. Initially, patients are divided into a training (175 patients) and testing (24 patients) sets. The preprocessing pipeline segments tumors and normalizes the contrast intensity. Features are retrieved from the segmented ROIs. After feature selection, relevant features are retained. Five machine learning models for survival prediction were examined. Three different experiment configurations (referred to throughout this chapter as Experiments/Results I, II and III) were defined for comparative performance assessment.

4.2.1 Data preprocessing

All the datasets include four MR weighted images (T1w, T2w, T2w-FLAIR, and post-T1w), which were first co-registered to 1mm^3 isotropic resolution and skull-

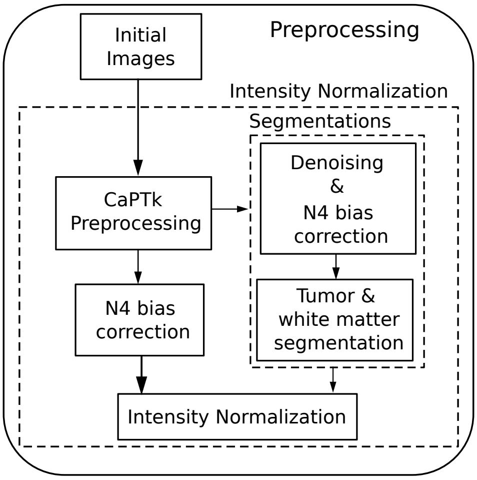


Figure 4.2: Preprocessing pipeline. Initial images are first co-registered and skull-stripped following CaPTk pipeline. Afterwards, the different contrast images are denoised and bias-corrected before obtaining the white matter and tumor segmentations. Finally, skull-stripped images are bias-corrected and intensity normalized using the segmentations.

stripped (Thakur et al., 2020) following *BraTS* preprocessing in CaPTk (Rathore et al., 2017). Then, denoising (Maggioni and Foi, 2012) followed by N4 bias correction (Tustison et al., 2010) was performed in order to obtain WM (Zhang et al., 2000) and tumor segmentations. The DL method nNUnet (Isensee et al., 2021) was utilized to segment the tumor into three distinct regions (i.e., enhancing tumor, non-enhancing tumor, and edema). These regions were employed for feature extraction in the radiomic system. On the other hand, N4 bias correction (Tustison et al., 2010) was applied on the skull-stripped images and these were next normalized dividing each by the mean intensity of the WM region contralateral to the tumor. This latter pipeline produces the images used as input to the radiomic system and the synthesis DL method. Figure 4.2 depicts the preprocessing pipeline.

4.2.2 Synthesis method by a self-supervised CNN

In Chapter 3 we have presented a joint Synthetic MRI approach for the computation of T1, T2, and PD parametric maps and the synthesis of different weighted images from only a pair of weighted inputs. A CNN trained mainly with synthetic data was employed. However, some synthesized weightings, such as T2w-FLAIR, presented relatively low quality presumably due to the exclusively synthetic training. Thus, if only a few weighted images are of interest, the issue referred to above can be overcome by extending the CNN into a self-supervised CNN to be trained with acquired weighted images with the desired contrast. Such an extension has been performed in this chapter and is graphically represented in Figure 4.3.

The original CNN was configured with two encoders for the two input weighted images, namely, T1w and T2w or T1w and T2w-FLAIR (in the case of synthesizing T2w-FLAIR or T2w, respectively). Then, the latent representations of each encoder

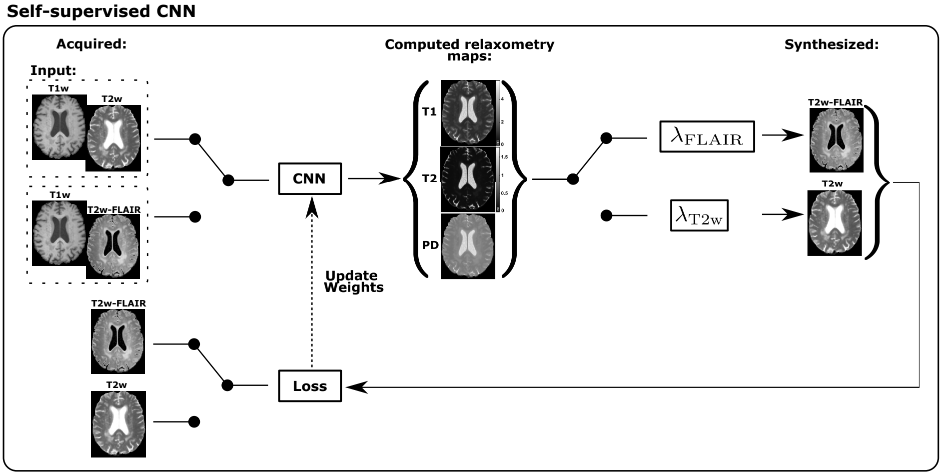


Figure 4.3: Overview of the self-supervised CNN. The input of the network are T1w and T2w for the synthesis of T2w-FLAIR, and T1w and T2w-FLAIR for the synthesis of T2w. Note that all the switches change depending of the weighting we want to synthesize. The lambda layers implement the theoretical pulse sequences equations that describe MR image intensity as a function of parametric maps and acquisition parameters.

were fused using a pixel-wise max function. Finally, we configured three decoders for the generation of the T1, T2, and PD parametric maps. In order to extend the CNN into a self-supervised CNN (see Figure 4.3) we have included a non-trainable lambda layer after the decoder's output. This lambda layer implements the theoretical pulse sequence equations that describe MR intensity of the output weighted images in relation with the aforementioned parametric maps and the acquisition parameters. These equations were previously explained in Subsection 2.1.4. Specifically, we employ Eq. 2.11 and Eq. 2.13 for the T2w and T2w-FLAIR, respectively.

The loss function used to train the self-supervised CNN, named L_{syn}, is computed in the weighted image domain as the average of the MAE between each acquired image and its synthesized counterpart. Specifically, let $m^k_{\text{acq}}(\mathbf{x})$ denote the intensity value of the $k-th$ acquired image at pixel $\mathbf{x}$, defined in some domain $\chi^k \subset \mathbb{R}^2$, and let $m^k_{\text{syn}}(\mathbf{x})$ be the synthesized image at that pixel location. Then:

$$\text{MAE}\left(\mathbf{m}^k_{\text{acq}}, \mathbf{m}^k_{\text{syn}}\right) = \frac{1}{|\chi^k|} \sum_{\mathbf{x} \in \chi^k} \left| m^k_{\text{acq}}(\mathbf{x}) - m^k_{\text{syn}}(\mathbf{x}) \right| \tag{4.1}$$

with $\mathbf{m}^k_{\text{acq}}$ and $\mathbf{m}^k_{\text{syn}}$ vectors that represent the image intensity values of the acquired and synthesized images respectively in all the pixels belonging to domain χ^k with cardinality $\left|\chi^k\right|$. Then, the loss function is defined as:

$$L = \frac{1}{M} \sum_{k=1}^{M} \text{MAE}\left(\mathbf{m}^k_{\text{acq}}, \mathbf{m}^k_{\text{syn}}\right), \tag{4.2}$$

with M the overall number of images entering the average (i.e. the batch size).

Multicontrast Gliblastoma, composed of a total of 24 patients acquisitions (see Section 1.4), was used to train the self-supervised CNN following a leave-one-out scheme (i.e., a total of 24 models were trained). For training each model, one patient was used for testing and the remaining 23 patients were randomly split between training (18 patients, representing approximately 80%) and early-stopping validation (5 patients, representing approximately 20%). The test patients of each model were then used for testing the radiomic system described next in Section 4.2.3. Hence, notice that neither the self-supervised CNN nor the radiomic system has never seen any tested patient along the entire pipeline.

During training, the loss function was optimized using the Adam optimization algorithm (Jais et al., 2019) with a learning rate of 1×10^{-4} and, as previously stated, early-stopping was used to avoid overfitting. Further, we empirically fixed the batch size to 32 images. We ran the code using the TensorFlow backend (Abadi et al., 2016) on a single NVIDIA GeForce GTX 1070. The total learning took approximately one hour of computation time for each model. Note that once the network has been trained, the computation time reduces to a few seconds.

4.2.3 Radiomic system

A radiomic system for survival prediction was trained to classify patients with long term survival rates (survival $>$ 480 days). The threshold of 480 days (i.e., 16 months) was chosen in order to achieve a balance between groups sizes in the test dataset. The datasets *BraTS2020*, *TCIA*, and *Dataset22* (175 patients in total, see Appendix C.1) were used to train the radiomic system. *Multicontrast Glioblastoma* (24 patients) was used for testing in coordination with the synthesis method previously described in Section 4.2.2.

Starting from a total of 117,088 handcrafted features extracted from the structural MR weighted images, detailed in Appendix C.2, we trained the radiomic system following a nested cross-validation scheme (outer = 5-fold; inner = 10-fold). Feature selection methods Ding and Peng (2005); Moore and White (2007) were repeated in each outer split to reduce the possible bias produced if training were done on a single cross-validation split. For each outer split, the model with the lowest Brier loss in the inner split was chosen. Note that five models were selected for the following screening due to the 5-fold decision of the outer split. Each of these models were then validated with the validation data corresponding to its outer split. Finally, the model with the best performance, measured with the AUC, was selected. The whole radiomics pipeline is outlined in Appendix C.2.

Using this methodology, the best model for each of the three scenarios described below was chosen:

1. When the four weighted images (i.e., T1w, T2w, T2w-FLAIR, and post-T1w) were used as input of the radiomic system, the selected model turned out to be a extreme gradient boosting (XGB) with 17 features, two of which belong to the T2w-FLAIR and another two to the T2w.

2. When only T1w, T2w and post-T1w images were used as input and the channel fed with T2w-FLAIR was discarded, the resulting model was a logistic regression (LR) classifier with 16 features.
3. When only T1w, T2w-FLAIR and post-T1w images were used as input and the channel fed with T2w was discarded, a support vector machine (SVM) classifier with 16 features was selected.

Hereinafter, these three models are termed XGB17, LR16, and SVM16, respectively. Features selected for each of the previous models are listed in Appendix C.3 Tables C.6, C.7, and C.8, respectively.

4.3 Experimental work

We carried out three test experiments in order to assess the possibility of using synthesized images as input of a radiomic system for survival prediction. In all of them the radiomic system was tested with *Multicontrast Glioblastoma*, with the provisions made in the previous section to avoid test data entering the process twice. These experiments are detailed next:

I) XGB17 was tested with the acquired T1w, T2w, T2w-FLAIR, and post-T1w images as inputs.

II) XGB17 was tested replacing one of the acquired inputs by its synthesized counterversion. T2w-FLAIR and T2w were considered for the replacement, one at a time. Therefore, in this experiment three of the inputs of the radiomic system were acquired and one was synthesized. As previously stated, these synthesized images were the test images from the leave-one-out of the synthesis method, so no overlap between the training and testing splits occurs in either the synthesis or in the radiomic system.

III) Models LR16 and SVM16 were tested without considering as input T2w-FLAIR and T2w, respectively. Note that the radiomic systems used in this third experiment, had been built with only three input channels.

Performance assessment is two-fold. On the one hand, we evaluated the quality of the synthesized images. In addition to visual assessment, we also carried out a quantitative analysis using the well-known measures MSE, SSIM, and PSNR as synthesis quality metrics (Moya-Sáez et al., 2021). These metrics have been defined within a 3D domain between both the synthesized and the acquired images, specifically, within the smallest cube that comprises the foreground of each volume); hence χ^k in Eq. (4.1) is the intersection between this domain and the k-th acquired image. Thereafter, the mean and standard deviation values across patients were calculated.

On the other hand, in order to compare the performance of the radiomic system with the different experiment configurations, we computed area under the curve (AUC), accuracy, precision, recall and F1-score as classifier performance metrics. These metrics were reported as the average value over the two classification classes. Additionally, for the sake of completeness, we analyzed the predicted probabilities of survival obtained at the output of the radiomic system for the pair of experiments I–II and I–III. To this end, we calculated the R^2 value, customarily used in linear regression, to measure how predicted probabilities of experiments II and III deviated from the identity function at abscissae equal to the probabilities of experiment I. The ICC (Koo and Li, 2016) between these pairs was also measured and boxplots of the probability differences for these pairs of experiments were constructed.

4.4 Results

Figure 4.4 shows a representative slice of the synthesized and the corresponding acquired weighted images for several test glioblastoma patients. Overall, the synthesized images are close to the acquired versions regarding the structural information and contrast between tissues in both healthy and pathological regions. Particularly, the contour and intensity of the different lesion areas are similar in both. However, it can be noticed that synthesized images tend to appear slightly blurrier than the acquired ones. In contrast, note that in Patient 3 the synthesized T2w-FLAIR image does not suffer from motion artifacts, which are indeed present in the acquired image.

Additionally, Table 4.1 shows the mean and standard deviation (SD) values computed across patients of the synthesis quality metrics (i.e., SSIM, MSE, and PSNR) between synthesized and acquired images for T2w-FLAIR and T2w. SSIM is a value ranging between 0 and 1, and the value 1 is only reachable for two identical images. The high values of PSNR and the low values of MSE show low error between the synthesized and the acquired weighted images for both the T2w-FLAIR and T2w. All the metrics improve considerably with respect to those obtained in Chapter 3 for the T2w-FLAIR weighted image. Note that the comparison of the values obtained for T2w are not representative since this weighted image was input to the CNN in that work.

Table 4.1: Synthesis quality metrics used to evaluate the capability to synthesize T2w-FLAIR and T2w weighted images. These metrics are the MSE, SSIM, and PSNR. Mean and standard deviation (between brackets) computed across patients are reported. The metrics were calculated between both the synthesized and the acquired images.

	MSE	**SSIM**	**PSNR**
T2w-FLAIR	0.0163 (0.0104)	0.7595 (0.0474)	23.7975 (2.0011)
T2w	0.0742 (0.0319)	0.7845 (0.0616)	25.8934 (1.9032)

Figure 4.5 shows the AUC, accuracy, precision, recall, and F1-score achieved for the radiomic system for the three different experiment configurations (experiments I, II, and III defined in section 4.3). All the metrics are substantially better in the case of using a synthesized image rather than using a system without such a weighted image for both T2w-FLAIR and T2w images. The comparison between using an acquired and a synthesized image shows that the performance of the system

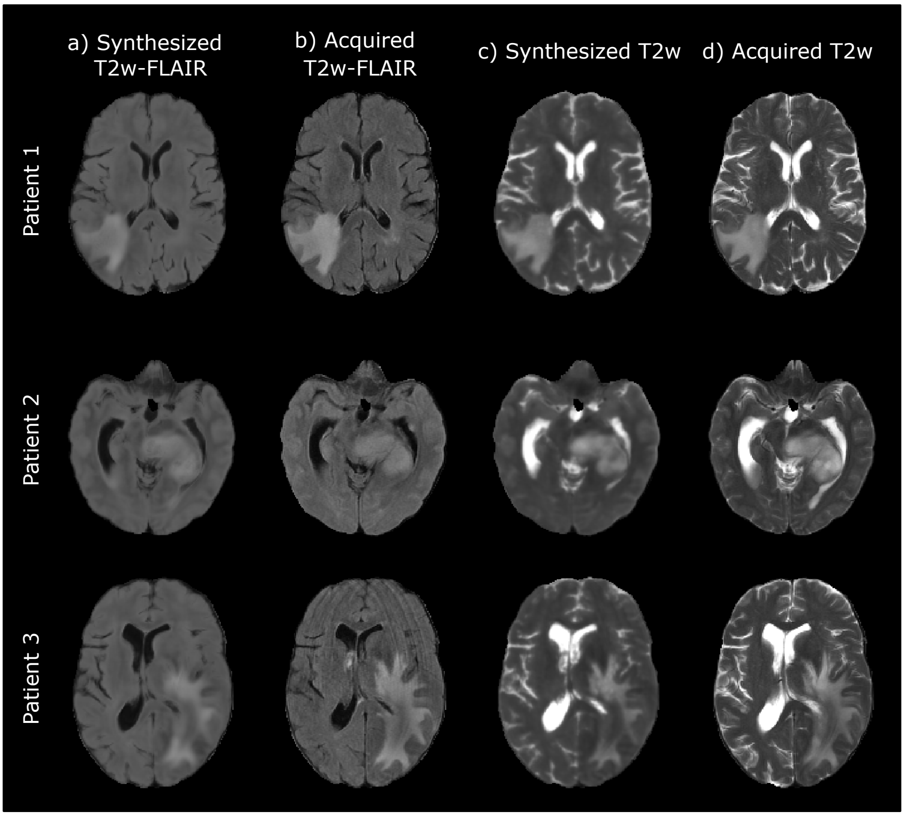


Figure 4.4: A representative axial slice of the synthesized images by the self-supervised CNN for different test patients of *Multicontrast Glioblastoma*. a) Synthesized T2w-FLAIR images. b) Corresponding actually acquired T2w-FLAIR images. c) Synthesized T2w images. d) Corresponding actually acquired T2w images.

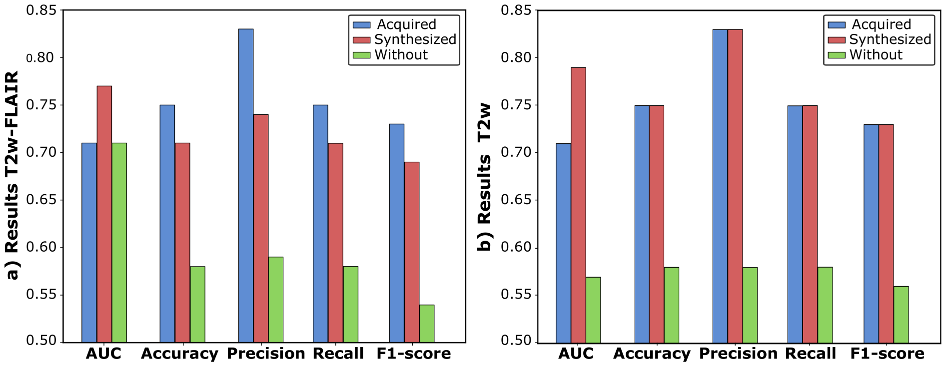


Figure 4.5: AUC, accuracy, precision, recall, and F1-score of the radiomic system tested with *Multicontrast Glioblastoma* when a) the T2w-FLAIR and b) the T2w input is I) acquired, II) synthesized, and III) when a radiomic system trained from scratch without considering any of these weighted images as input is used instead.

does not diminish in terms of AUC, and only suffers from a slight degradation in terms of the other performance metrics when synthesizing the T2w-FLAIR. Such degradation is not observed with the synthesized T2w.

Figure 4.6 shows scatter plots of the output predicted probabilities of experiments I–II and I–III. The ground-truth labels and labeled predictions (i.e., survival $>$ 480 days or not) are also displayed. Additionally, R^2 values from the identity linear regressions are provided. A better agreement of points in plots in the upper row (experiments I–II) compared to the plots in the lower row (experiments I–III) can be observed. This better agreement is also confirmed with the higher values of R^2 obtained. Additionally, the ICC values measured are 0.983 and 0.292 for T2w-FLAIR and 0.964 and 0.027 for T2w, for the pair of experiments I–II and I–III, respectively. These ICC values correspond with an excellent agreement according to Koo and Li (2016).

Finally, Figure 4.7 shows boxplots of the probability differences for the pair of experiments I–II and I–III, for both T2w-FLAIR and T2w images. As can be seen, the median of the boxplots in the pair I–II is closer to zero than in the pair I–III for both the T2w-FLAIR and the T2w. A lower interquartile range (IQR) in boxplots of experiments I–II compared to experiments I–III can be also observed, together with a median shift from zero in the I–III experiment, an effect which is more prominent for T2w.

4.5 Discussion

In this chapter, we have thoroughly analyzed the replacement of an actual acquired weighted image with its synthesized counterversion for predicting survival of glioblastoma patients with a completely independent radiomic system. Starting from two acquired weighted images, we synthesized a new weighted image using a self-supervised CNN-based method. The radiomic system was trained using as

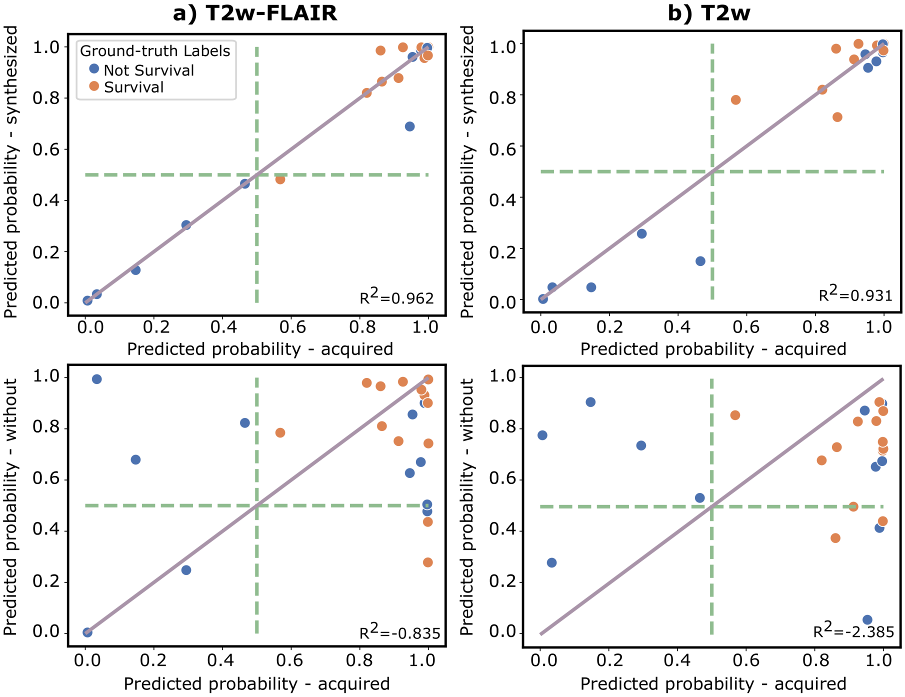


Figure 4.6: Scatter plots of the predicted probabilities obtained at the output of the radiomic system for the experiment with the acquired versus the synthesized images (top row) and versus a radiomic system trained from scratch without considering this weighted image as input (bottom row). The plots are shown for a) T2w-FLAIR and b) T2w. Dashed lines represent the threshold fixed in the radiomic system to classify survival (>480 days). Each point represents each glioblastoma patient and its color corresponds to the ground-truth labels. R^2 values from the identity linear regressions are provided. Note the good agreement of points in upper plots compared to lower plots.

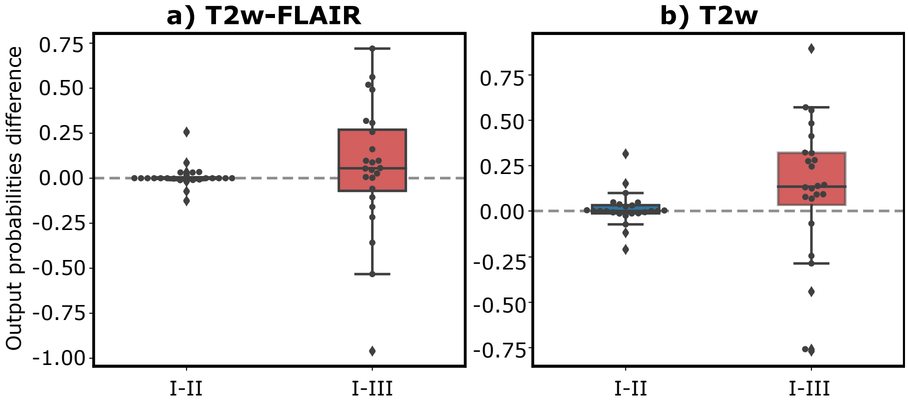


Figure 4.7: Boxplots of the differences between the probabilities obtained at the output of the radiomic system for the experiments with the acquired versus the synthesized images and versus a radiomic system trained from scratch without considering this weighted image as input (i.e., pair of experiments I–II and I–III, respectively). The boxplots are shown for a) T2w-FLAIR and b) T2w. Each point represents the probability difference for each glioblastoma patient and dashed line correspond to zero difference. Note that the median of the boxplots is closer to zero in the pair I–II compared to I–III.

input acquired images only. Then, the system was tested using as input acquired images, on one side, and replacing one acquired image with a synthetic image, on the other. We also compared performance with a system trained from scratch ignoring this additional channel. T2w-FLAIR and T2w contrasts were used for replacement.

The self-supervised approach proposed in this work allows us training the DL method without the need of reference parametric maps; thus, only weighted images are necessary. In addition, the incorporation of physical knowledge in the training by means of the theoretical pulse sequence equations could leverage the quality of the synthesized images and speed up the training process.

Results show that multicontrast-demanding quantitative applications, such as Radiomics, can be leveraged by synthesized images. Synthesized images may allow widespread usage of these radiomic systems in clinical practice, by retrospectively completing databases with missing modalities and/or replacing artifacted images. An example of the latter can be found in Figure 4.4, where the synthesized T2w-FLAIR image of Patient 3 does not suffer from motion artifacts, which are indeed present in the acquired image. Further, these synthesized images have the potential to speed up acquisition protocols by replacing some acquired images with their synthesized counterversions. Particularly, removing T2w-FLAIR or T2w from an average brain protocol may reduce the overall scan duration on the order of 20% and both of them are artifact-prone sequences due to their sensitivity to motion. Thus, our results allow us to state that Synthetic MRI does add up to glioblastoma survival prediction within a Radiomics-based approach. Indeed, the network described in Chapter 3 is prepared to synthesize more than one contrast, so for protocols with more sequences than those used in this paper, higher reductions could be potentially achieved.

It has been shown in this work that synthesized weighted images are visually similar to the acquired images in both healthy and pathological tissues. However, it can be observed that synthesized images tend to exhibit a slight blurriness compared to the acquired ones. This occurrence is likely due to an implicit filtering effect introduced by the CNN during the self-supervised learning process. Additionally, the values of image quality metrics prove the agreement between the synthesized and the actually acquired images. The performance achieved with the synthesized images in the radiomic system is not only close to the performance achieved using the acquired images, but also substantially better than using a model trained without that weighting. This is confirmed by the classifier performance metrics (i.e., AUC, accuracy, precision, recall and F1-score) for both T2w-FLAIR and T2w. The R^2 of the identity linear regression and the ICC values also support this finding. Note that an ICC value above 0.9 is considered excellent (Koo and Li, 2016). It is worth noting that the synthesized images input to the XGB17 model improve the AUC compared to acquired images. This might be caused by the implicit filtering undergone during the synthesis procedure. Moreover, accuracy values obtained from the radiomic system are in par with other radiomic systems which rely exclusively on acquired images (Tewarie et al., 2021), and set the survival threshold for classification, similarly as we do, to achieve balance between groups sizes.

This work has several limitations. The test experiments were carried out on *Multicontrast Glioblastoma*, composed of a cohort of 24 glioblastoma patients. We made this design decision because *Multicontrast Glioblastoma* is the only dataset in which the pulse sequence and the acquisition parameters remained steady across patients, and the self-supervised proposed synthesis method depends on these parameters. This dependence has the advantage of making the process specific to this parameter setting, so higher synthesis quality can be expected. The downside is the inherent limitation to this particular setting. Nevertheless, the self-supervised method can be easily extended to accommodate more parameter values for which acquisitions are available. In addition, experiments on a larger cohort would be advisable to further support our conclusions. Further, a multi-institutional study could be necessary to analyze the capability of the system to generalize.

One might argue about the need to synthesize weighted images to feed a radiomic system since parametric maps were already available. Note that these maps were generated as a previous step to synthesize precisely the weighted images. Certainly, the radiomic system could have been designed to deal directly with these maps, and this is a topic in which some other predictions have been properly proposed (Pirkl et al., 2021). However, we have two reasons that support our design choice. First, the glioblastoma datasets we have used for training the radiomic system do not include parametric maps. Second, the approach proposed in Chapter 3 has been trained with glioblastoma-free images, both from synthetic data as well as a small dataset of actual acquisitions. Hence, we have extended our original method with a self-supervised procedure to include glioblastoma information in our pipeline.

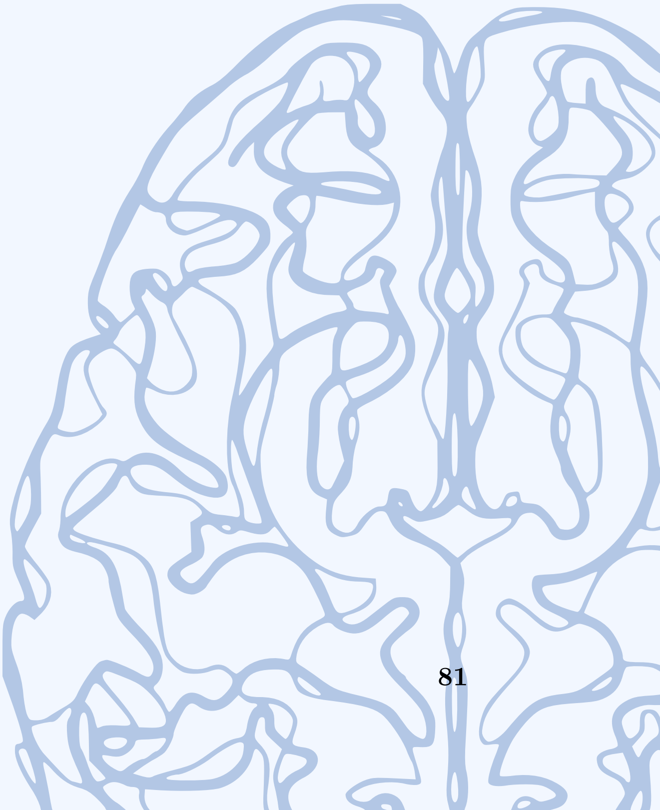
As future work, performing the experiments synthesizing the post-contrast weighted images might be of interest from a clinical point of view, in order to avoid the administration of contrast agents to patients. Recently, some promising works (Dai et al., 2020; Kleesiek et al., 2019) that attempt to carry out the translation between pre-contrast and post-contrast weighted images have been proposed, but they still have some limitations. On the one hand, the lack of versatility since neither the pulse sequence nor the acquisition parameters can be controlled in these methodologies. On the other hand, the qualitative nature of weighted images makes the replacement of contrast agents in these methods less explicable since they are not based on parametric maps.

4.6 Conclusions

In conclusion, in this chapter we assessed the performance of a radiomic system when an input actually acquired was replaced with a synthesized counterversion. To this end, we synthesized realistic T2w-FLAIR and T2w images in a glioblastoma dataset with a deep learning approach. Furthermore, a radiomic system for survival prediction, which can classify patients in two groups (survival $>$ 480 days and $\leq$ 480 days) was built. We evaluated the effects of the synthesized weighted images in the radiomic system performance. Results support the utility of using synthesized images to feed a radiomic system for survival prediction of glioblastoma patients.

Part IV

Synthetic MRI contributes to the detection of tumor enhancement



5

Brain tumor T1w-enhancement prediction from pre-contrast conventional weighted images using synthetic multiparametric mapping and deep learning

Contents

The work in this chapter is currently in review as:

- **Moya-Sáez, E.**, de Luis-García, R., Nunez-Gónzalez, L., Alberola-López, C., and Hernández-Tamames, J.A. Brain tumor T1w-enhancement prediction from pre-contrast conventional weighted images using synthetic multiparametric mapping and deep learning. *Submitted.*

5.1 Purpose

The assessment of BBB breakdown in brain tumors is usually performed by the acquisition of a T1w image after the injection of a GBCA. Thus, after the injection, the T1w and post-T1w are visually compared in order to discover extravasation of GBCA into the perivascular space (Hattingen et al., 2017). As we introduced in Chapter 1, both the usage of GBCAs and the assessment based on visual inspection present associated problems.

Parametric maps (i.e., T1, T2, and PD) could be a potencial solution due to their ability to facilitate quantification of subtle changes within the tissues (Badve et al., 2017; Hattingen et al., 2017; Lescher et al., 2015; Nunez-Gonzalez et al., 2022). Hence, parametric maps could be key for predicting BBB damage without GBCAs, as shown by Nunez-Gonzalez et al. (2022). This study differentiates between healthy and abnormal tissue and, particularly, tissue with and without T1w-enhancement using only pre-contrast maps (Nunez-Gonzalez et al., 2022).

The main issue that limits the applicability of this method is that maps are not commonly acquired in clinical practice due to their lengthy relaxometry acquisitions. Nowadays, fast multiparametric mapping techniques, such as MR Fingerprinting (Ma et al., 2013) or MAGiC (Warntjes et al., 2008), have taken the stage. These techniques, albeit faster, still have time limitations; moreover, these sequences are scarcely available worldwide. However, an alternative solution could be the computation of parametric maps from conventional weighted images using DL as we have proposed in Chapter 3. In that chapter the computed maps for healthy volunteers were compared with those obtained with traditional sequences, but their diagnostic value was not clinically validated. Specifically, no validation was performed focused on T1w-enhancement prediction in oncological patients.

In this chapter, we propose to predict post-contrast T1w-enhancement in glioma patients from pre-contrast conventional weighted images using DL-computed parametric maps.

5.2 Methods

Three different datasets were employed in this work; one of them (*GLIOMA*) for training and testing the DL approach, and the other two (*UPenn-GBM* and *Relaxometry Brain-MRI*) for testing the generalization capabilities of the DL method. These datasets have been previously described in Chapter 1. All the datasets include T1w, T2w, T2w-FLAIR, and post-T1w images. In addition, *GLIOMA* and *Relaxometry Brain-MRI* include T1, T2, and PD parametric maps acquired with MAGiC* (Warntjes et al., 2008) and relaxometry sequences, respectively. An overview of the proposed DL method together with the experiments can be found in Figure 5.1. Details about the the number of participants selected for each dataset are also shown.

---

*Commercial implementation of the QRAPMASTER technique for quantitative imaging.

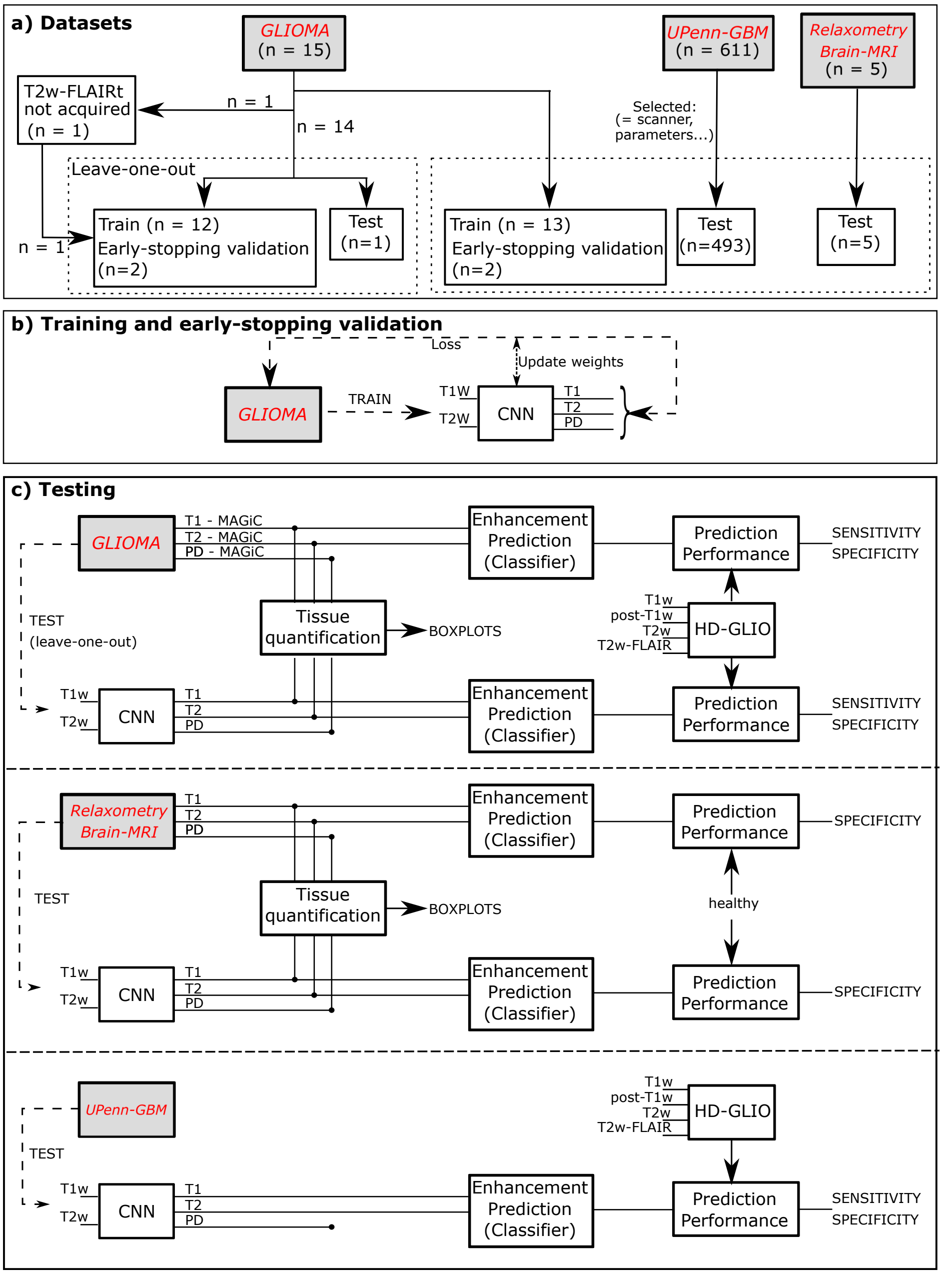


Figure 5.1: Flow diagram. a) Participants selection of each dataset, namely *GLIOMA*, *UPenn-GBM*, and *Relaxometry Brain-MRI*, and training/early-stopping validation/test sets splitting. b) Pipeline of the proposed approach for training the DL method with the *GLIOMA* dataset. c) Testing and experimentation employed for each dataset depends on the data available in each of them.

5.2.1 Data processing

First, all weighted images were reoriented to match the orientation of the standard MNI152 (Evans et al., 2012) with FSL (Jenkinson et al., 2012). After that, images were skull-stripped using HD-BET (Schell et al., 2019) followed by a linear registration to the T1w using FLIRT tool of FSL (Jenkinson et al., 2012). Next, the non-skull-stripped images were registered using the previous linear transformation and brain masks were applied to all registered images in order to null non-brain voxels. The last preprocessing step is the normalization of the weighted images by dividing each of them by its average intensity excluding the background. The segmentation of WM, and GM tissues were obtained with FAST tool of FSL (Jenkinson et al., 2012). In addition, different regions were defined: T1-weighted contrast-enhancement (T1e), non-enhancing T2-weighted hyperintensity (T2h) — jointly referred to as abnormal tissue (ABN) —, and normal white matter (nWM), which corresponds to the WM segmentation.

5.2.2 T1w-enhancement assessment

Different methods have been recently proposed for the automatic classification of voxels with and without T1w-enhancement. We will use two of them which follow different approaches; the first one needs GBCA but the second one does not. Next, we detail the characteristics of both and how we have employed them in this work.

HD-GLIO It is a DL segmentation tool (Kickingereder et al., 2019), which uses pre-contrast (T1w, T2w, T2w-FLAIR) and post-contrast (post-T1w) images as inputs. It segments voxels with T1e, and voxels with T2h. It is important to highlight that this tool was trained with annotations performed by experts. In this work, *GLIOMA* and *UPenn-GBM* datasets were input to the HD-GLIO for obtaining ground-truth segmentations.

Voxel-wise statistical prediction Nunez-Gonzalez et al. (2022) recently proposed a method for voxel-wise classification of normal/abnormal tissue and T1w-enhancement/non-enhancement based only on pre-contrast (i.e., without GBCAs) T1 and T2 maps obtained with MAGiC. Further details on this method can be found in Appendix D. Following this methodology, in this work we considered two classification problems: classification-I (C-I) ABN versus nWM, and classification-II (C-II) T1e versus the union of nWM and T2h. For each classification problem, we selected the metric that showed the best performance in Nunez-Gonzalez et al. (2022) and its optimal operating point, i.e., normlog with threshold 8.44 for C-I, and normT1T2 with threshold 1344 ms for C-II.

5.2.3 Synthesis of parametric maps

The computation of parametric maps is performed with a CNN that extracts the quantitative T1, T2, and PD information embedded in the weighted images (see Figure 5.1.b). The CNN was configured with two encoders —one per input (i.e., the T1w and the T2w)—, and three decoders —one per desired parametric map— The pipeline has three steps: 1) Each weighted image is input to its encoder. 2) The output of the encoders are fused with a pixel-wise max function into a shared representation. 3) The shared representation is input to each decoder. The network processes the volumes slice-wise.

All encoders share the same architecture, which is inspired by the UNet to exploit information at larger spatial scales (Ronneberger et al., 2015). Skip connections were used to avoid the loss of details induced in the downsampling path. Also the decoders share the same architecture; in this case a fully-convolutional network is employed. More details about the architecture can be found in Section 2.5.4.

Both encoders and decoders were trained together. Training was supervised with the loss function:

$$L = ||(\mathrm{T1}_c - \mathrm{T1}_{\mathrm{GT}})||_{\ell 1} + ||(\mathrm{T2}_c - \mathrm{T2}_{\mathrm{GT}})||_{\ell 1} + ||(\mathrm{PD}_c - \mathrm{PD}_{\mathrm{GT}})||_{\ell 1}, \qquad (5.1)$$

where the subscripts c and GT refer to the DL-computed and the ground-truth (i.e., MAGiC) maps, respectively. The $l1$-norm was chosen to be robust against misregistration of input images. The loss function was minimized using Adam (Jais et al., 2019) with a learning rate of $1e^{-4}$. The batch size was empirically set to 4 slices. The number of epochs was determined in execution time by early-stopping (10 epochs with a loss reduction less than 0.001 in early-stopping validation set) to avoid underfitting and overfitting. Pre-training with synthetic data is performed following the steps detailed in Chapter 3, in order to both speed up the training process and obtain good performance with a relatively small training set. We implemented our model with Tensorflow v.2.4.0.

5.2.4 Training, early-stopping validation and test sets

First, the DL method was trained following a leave-one-out scheme with *GLIOMA*; hence, one patient was left aside for testing and the remaining patients were randomly split into training (12 patients) and early-stopping validation (two patients). Note that the patient without T2w-FLAIR was not included in any test set because HD-GLIO can not be executed on it, but it was additionally included in the training/early-stopping validation set as shown in Figure 5.1.a). Thus, a total of 14 data splits were performed. Patients may exhibit T1w-enhancement or not. To ensure a balanced representation of this condition in the early stopping validation set, we took care to include one patient with this condition and one without.

Second, to assess generalization capability, we performed a new training, but, in this case, with the whole *GLIOMA* dataset, i.e., no patient was left aside for testing. Specifically, we used one random split of *GLIOMA* in training (13 patients) and early-stopping validation (two patients) sets. Testing was carried out on two separate datasets (*UPenn-GBM* and *Relaxometry Brain-MRI*).

5.3 Experimental work

Two kind of experiments have been conducted on the test sets for the validation of the proposed approach. These experiments are labeled in Figure 5.1.c) as *Tissue quantification* and *Prediction performance*. The former is intended to evaluate similarity between the DL-computed and the acquired maps. The latter evaluates the capability of DL-computed maps for predicting ABN and T1e tissues. We stress that the validation of the proposed approach does not correspond with the early-stopping validation, which is performed during the network training.

Specifically, and following Figure 5.1.c), for *GLIOMA* the validation of the proposed approach was performed using the test patient for each trained model in the leave-one-out. As for the *tissue quantification* of T1, T2, and PD parameters, MAGiC and DL-computed parametric maps were visually compared and the percentage error was voxel-wise computed as the difference between the two maps (DL-computed and acquired) normalized by the acquired MAGiC map. The WM and GM segmentations for each test patient were applied to both types of parametric maps. Then, boxplots of the voxel-wise percentage error between the DL-computed and acquired MAGiC maps for the WM and GM tissues were constructed. For the second experiment, classifications were carried out with both acquired MAGiC and our DL-computed parametric maps for *GLIOMA*, and their results were compared in terms of sensitivity and specificity. These two parameters were computed for each classification problem performed with the metrics and thresholds previously stated in section 5.2.2. The reference segmentation for computing sensitivity and specificity was obtained with HD-GLIO from the weighted images (see Figure 5.1.c). Statistical differences were calculated by means of a significance test.

The dashed horizontal lines in Figure 5.1.c) depict the tests carried out on the two remaining datasets (*Relaxometry Brain-MRI* and *Upenn-GBM*). These tests allow us to assess the generalization capability of our method since they are conducted on datasets that were not used for training. The *tissue quantification* of T1, T2, and PD parameters was also conducted on *Relaxometry Brain-MRI*; in this case, WM and GM tissue values were compared between our synthesized maps and the maps included in *Relaxometry Brain-MRI* by means of the voxel-wise difference between the two maps normalized by their voxel-wise mean. As for the second experiment (labeled as *Prediction performance*), it was conducted on both datasets. As for *Relaxometry Brain-MRI* only the specificity on the classification was calculated since the dataset consists of healthy volunteers. Differences are also tested by a significance statistical test. For *UPenn-GBM*, voxel-wise classifications were performed for the DL-computed maps and, afterwards, sensitivity and specificity were computed using HD-GLIO as the reference segmentation.

For the three datasets, the normality of the samples of sensitivity and specificity are tested with a Shapiro test. For normal data, values are reported as mean and SD. Otherwise, values are reported as median and IQR. As for statistical tests, either the paired t-test or Wilcoxon signed rank test have been used according to the normality of the samples. P-values are reported for an unilateral test (i.e. the alternative hypothesis is “the value obtained for our maps are greater than for the other method maps”).

Table 5.1: Comparison of the classification performance with MAGiC and DL-computed maps for test patients of *GLIOMA*. Median across patients and IQR are reported. Bold text represents significant differences (p-value < 0.05) between DL-computed and MAGiC maps.

	(C-I) ABN vs. nWM		**(C-II)** T1e vs. non T1e	
	Voxel-wise MAGIC	**DL-computed**	**Voxel-wise MAGIC**	**DL-computed**
Sensitivity	89.35% (IQR: 8.93%)	88.37% (IQR: 17.30%)	87.29% (IQR:16.62%)	93.26% (IQR: 14.05%)
	p-value = 0.4029		**p-value = 0.0015**	
Specificity	93.61% (IQR: 2.14%)	95.21% (IQR: 1.13%)	94.95% (IQR: 2.88%)	95.59% (IQR: 1.81%)
	p-value = 0.0011		**p-value = 0.0054**	

5.4 Results

5.4.1 DL-computed parametric maps in *GLIOMA* dataset

Figure 5.2 shows, for a representative test patient, both the acquired MAGiC maps and the DL-computed maps. The voxel-wise percentage error between each pair of maps is also shown. It can be noticed that the DL-computed maps exhibit more blurring, with most of the differences located at the interfaces between tissues. This effect might be caused by partial volume effects and misregistration of the input weighted images. A comparison of tissue values of the WM and GM between the DL-computed and MAGiC maps is shown in Figure 5.3.

5.4.2 T1w-enhancement prediction in *GLIOMA*

Table 5.1 shows shows the sensitivity and specificity values obtained for each classification (i.e., C-I and C-II defined in Section) for both acquired MAGiC and DL-computed maps. Also for both cases, the segmented T1e and T2h regions obtained through the respective classifications are shown in Figure 5.4 for a representative test patient. The segmentations obtained with HD-GLIO from weighted images are also shown as reference.

5.4.3 Results on *Relaxometry Brain-MRI* and *UPenn-GBM*

Boxplots of the WM and GM percentage error between the DL-computed maps and those in *Relaxometry Brain-MRI* are shown in Figure 5.5. On the other hand, the voxel-wise classifications applied to the DL-computed maps reported mean specificities of 95.83% (SD: 0.94%) and 98.83% (SD: 0.16%) for C-I and C-II, respectively, while for the relaxometry maps from *Relaxometry Brain-MRI* the specificity lowered to 82.08% (SD: 5.93%) for C-I and 83.36% (SD: 6.03%) for C-II. Significant differences were found in both cases ($p<0.01$).

Figure 5.6 shows a comparison between the segmentation obtained with the voxel-wise classification from DL-computed maps and the reference HD-GLIO segmentation for the *UPenn-GBM* dataset. The sensitivities and specificities for the voxel-wise classifications with the DL-computed maps are shown in Table 5.2.

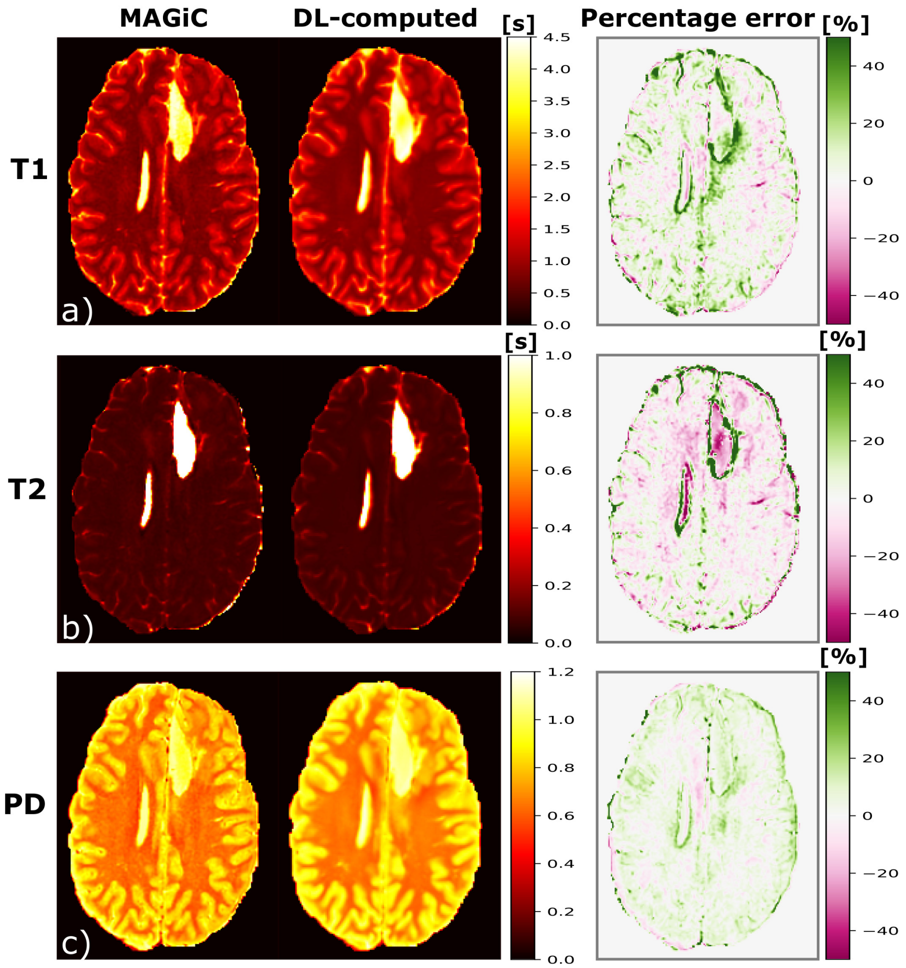


Figure 5.2: A representative axial slice of MAGiC and corresponding DL-computed parametric maps for a test patient of *GLIOMA*, who presents T1w-enhanced tissues. The voxel-wise percentage error is also represented for each pair. a) T1 map, b) T2 map, and c) PD map.

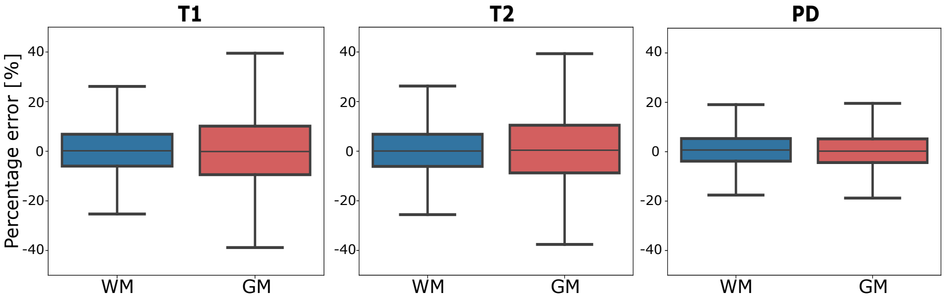


Figure 5.3: Boxplots of the WM and GM voxel-wise percentage error between the DL-computed and the MAGiC parametric maps for test patients of the *GLIOMA* dataset.

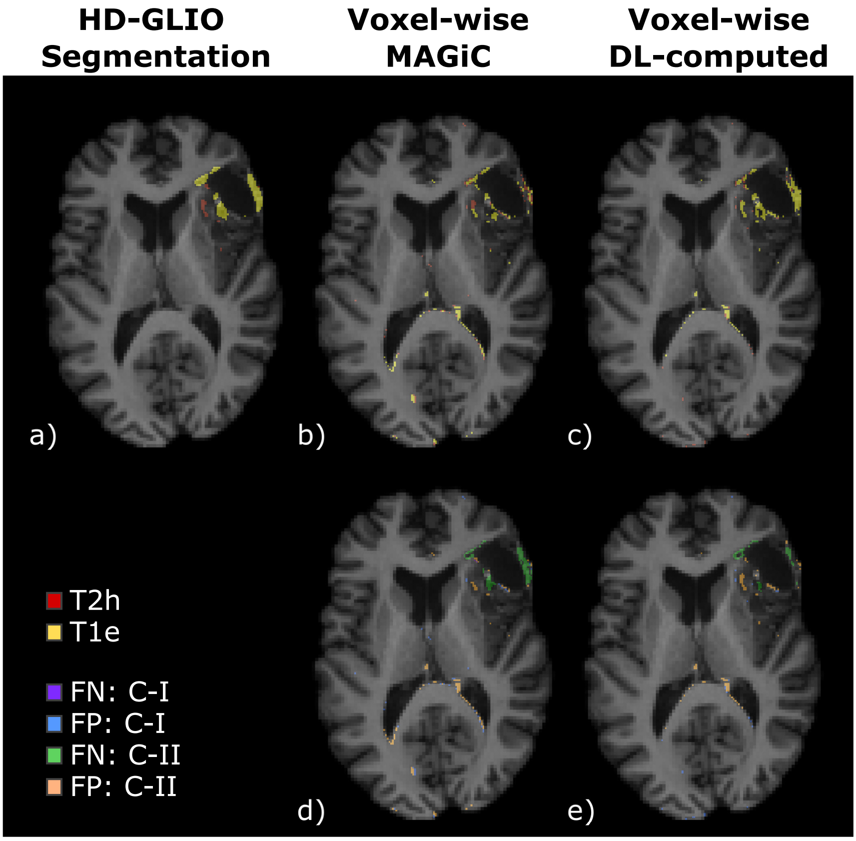


Figure 5.4: A representative axial slice of the T1e and T2h segmentations overlaid on the T1w images for a representative test patient of the *GLIOMA* dataset. a) Reference segmentation obtained with HD-GLIO from the four weighted images (T1w, T2w, T2w-FLAIR, and post-T1w). b). Segmentations obtained through the voxel-wise classifications from MAGiC parametric maps. c) Segmentations obtained through the voxel-wise classifications performed from DL-computed parametric maps. d) Errors of b) compared to a). e) Errors of c) compared to a). FN: false negative, FP: false positive.

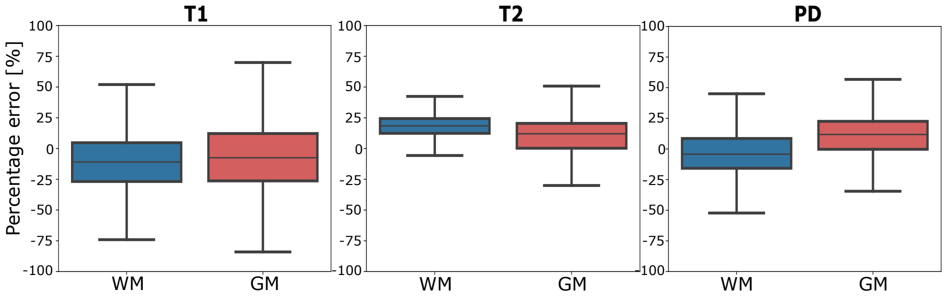


Figure 5.5: Boxplots of the WM and GM voxel-wise percentage error between the DL-computed and the relaxometry parametric maps for the *Relaxometry Brain-MRI* dataset.

Table 5.2: Classification performance of DL-computed maps for *UPenn-GBM* patients in terms of sensitivity and specificity. Median across patients and IQR for each classification problem are reported.

	(C-I) ABN vs. nWM	**(C-II)** T1e vs. non T1e
Sensitivity	91.23% (IQR: 10.14%)	81.04% (IQR: 24.21%)
Specificity	90.24% (IQR: 3.55%)	91.49% (IQR: 6.14%)

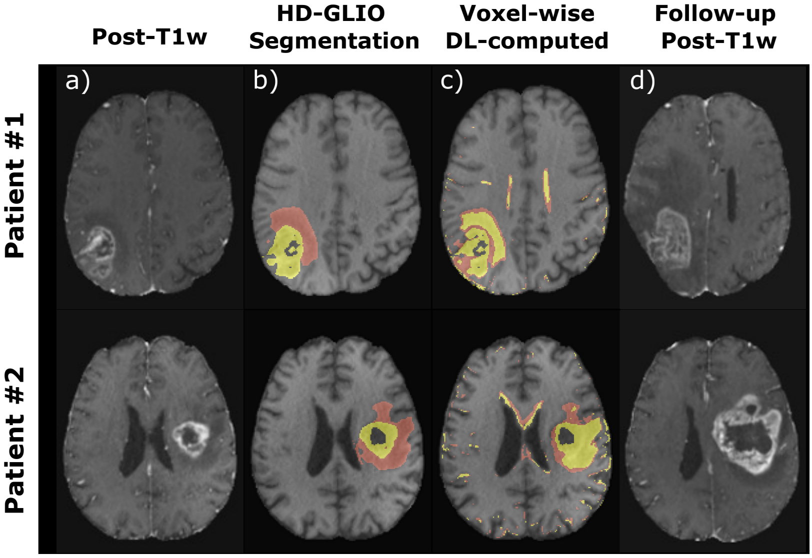


Figure 5.6: A representative axial slice of the T1e and T2h segmentations overlaid on the T1w images for different patients of *UPenn-GBM*. a) Post-T1w image. b) Reference segmentation obtained with HD-GLIO from the four weighted images (T1w,T2w, T2w-FLAIR, and post-T1w). c) Segmentations obtained through the voxel-wise classifications performed from DL-computed parametric maps. d) Post-T1w image from the follow-up acquisition. Red label represents T2h, whereas the yellow label represents T1e.

5.5 Discussion

In this chapter, we have proposed a method that enables the prediction of brain tumor T1w-enhancement from pre-contrast conventional weighted images through the DL computation of parametric maps. Supervised training was utilized to approximate the MAGiC parametric maps. We have validated the ability of DL-computed maps to predict abnormal and T1w-enhanced tissues without GBCAs and, also, their tissue quantification accuracy. Generalization capabilities of the proposed method have been also tested with multi-site, multi-vendor acquisitions.

Results have shown that the DL-based parametric mapping from only two conventional weighted images show tissue values similar to those obtained with relaxometry sequences or a fast multiparametric technique such as MAGiC. In addition, the performance of these DL-computed maps to discriminate normal/abnormal and T1w-enhanced/non-enhanced tissue is comparable to the performance obtained with MAGiC maps on the *GLIOMA* dataset. Both results could be replicated by testing the DL-based method with two additional datasets (*Relaxometry Brain-MRI* and *UPenn-GBM*) which were unseen during the training process. All datasets were collected in different centers and with 3T scanners of multiple vendors. It is worth noticing that neither inhomogeneity nor motion corrections were applied to the estimation of the *Relaxometry Brain-MRI* dataset. This could have an impact on the better performance of the DL-computed maps in terms of specificity in the voxel-wise classifications.

Interestingly, voxel-wise classifications show a slight overestimation of the T1e region compared to the HD-GLIO segmentation, but with some coincidence with the follow-up images shown in Figure 5.6. This region presents T1 and T2 values higher than those of the normal white matter, resulting in the region being classified as T1e. Nevertheless, we hypothesize that these higher values could be explained by an altered interstitial fluid mobility and increased water content in the perivascular space (Wardlaw et al., 2015), even when the BBB is not completely disrupted for gadolinium to pass through. Thus, this overestimated region could represent a BBB vulnerability and, therefore, it might be an appropriate follow-up biomarker, which is in agreement with recent reported findings (Lescher et al., 2015), and, apparently, with our results on the *UPenn-GBM* follow-up images (see Figure 5.6). However, further research is necessary to confirm this hypothesis.

The prediction of the T1w-enhancement region from pre-contrast parametric maps could bypass the injection of GBCAs, avoiding their related issues, as shown in Nunez-Gonzalez et al. (2022). However, the acquisition of parametric maps makes protocols considerably longer. Fast multiparametric mapping techniques as MAGiC do not fully address this issue since its acquisition is still not less than 4 min for a resolution of 1 x 1 x 3 mm. Thus, the computation of parametric maps from conventional weighted images, as proposed in this work, would favor the spread of this clinical application and, in a broader sense, the usage of parametric maps in clinical practice. Moreover, this approach proves valuable for retrospective analysis, enabling the computation of parametric maps for pre-existing and/or multicenter datasets, thereby empowering longitudinal or population studies.

If it is feasible to derive parametric maps from the weighted images and subsequently obtain T1w-enhancement measures from these maps, it follows that estimating tumor enhancement directly from the weighted images should also be viable. However, we choose to use the parametric maps as an intermediate step for several reasons: a) gaining resilience to variations in scanners and acquisition parameters by extracting the absolute scale parametric maps, b) offering information closer to the pathophysiology, and c) grounding our method on previous studies (Hattingen et al., 2013; Lescher et al., 2015; Müller et al., 2017; Nunez-Gonzalez et al., 2022), which have shown that quantitative T1 and T2 values support the rationale of T1w-enhancement prediction when replacing GBCAs, so that the results obtained with DL can be explained from a clinical viewpoint.

This work has several limitations. Acquisition time limitations lead to a mismatch in spatial resolution between the weighted images and MAGiC maps of the *GLIOMA* dataset. Consequently, the co-registration of all modalities is needed, which might induce misregistration artifacts and partial volume effects. Specifically, misregistration of the T1w and T2w images might be one of the main sources of errors of the proposed DL method despite the l1-norm used in the loss function. This can explain why most of the differences between the DL-computed and MAGiC maps are located at the edges (see Figure 5.2). Alternatively, the network might be inducing some blurring in the DL-computed maps, which could also explain these edge differences. Moreover, the T1w and T2w input images have varying acquisition parameters across datasets due to different institutional protocols. Although this study has indeed the value of showing comparable results when using other datasets for testing, these input parameter variations can potentially introduce errors in the computation of the parametric maps.

To mitigate the impact of variation in acquisition parameters, future work aims to incorporate them as additional inputs to the network. This approach would enable training with a more extensive and diverse dataset, potentially enhancing the model generalization capabilities. Furthermore, the extension of the approach to other clinical applications for diagnosing different pathologies is also planned (Deoni, 2010).

5.6 Conclusions

In conclusion, in this chapter we showed a proof-of-concept of predicting T1w-enhancement from pre-contrast conventional weighted images through the DL computation of parametric maps. The results suggest that these DL-computed maps might eventually have the potential to replace GBCAs for tumor T1w-enhancement prediction, without compromising performance and at no additional acquisition time.

6

Pre- and post-contrast simultaneous parametric mapping of glioblastomas from routine images for quantitative enhancement assessment

Contents

The work in this chapter has been published in:

- **Moya-Sáez, E.**, de Luis-Garcia, R., Hernández-Tamames, J.A., and Alberola-López, C. Pre and Post contrast Simultaneous Parametric Mapping of Glioblastomas from routine T1 weighted images for Quantitative Enhancement Assessment. *2023 ISMRM & SMRT Annual Meeting & Exhibition*, Toronto, Canada, June 2023; 2352.

6.1 Purpose

Throughout this thesis we have described the importance of evaluating BBB disruption in brain tumor assessment. In clinical practice this evaluation is typically performed based on the visual comparison between the T1w and post-T1w images. However, the qualitative nature and arbitrary scale of weighted image intensities hinder the usage of quantitative diagnostic methods, which could be relevant for distinguishing different post-treatment conditions, understanding biological changes within the tumor and minimizing interpretation errors of treatment effects.

Parametric maps are known to have a quantitative absolute scale, to be more robust against scanner imperfections and to be able to highlight subtle changes quantitatively, which are relevant ingredients to be qualified as biomarkers of impaired BBB (Blystad et al., 2017; Nunez-Gonzalez et al., 2022; Pirkl et al., 2021). In previous Chapter 5 we have shown the application of pre-contrast parametric maps to predict tumor enhancement without the usage of GBCAs. Alternatively, as we propose in current chapter, parametric maps could also complement the usage of GBCAs by enabling an automatic quantification of tumor enhancement. This might overcome the limitations associated with the standard visual assessment of tumor enhancement.

In Chapter 4 we have proposed a self-supervised Synthetic MRI approach for the computation of T1, T2, and PD parametric maps from only conventional weighted images. Self-supervised learning allowed us to compute reliable maps without the need of the corresponding reference maps for network training; in contrast, the training could be performed only with weighted images. Thus, in this chapter we propose the extension of that method for the computation of both pre- and post-contrast parametric maps. As we show, the computation of pre- and post-contrast maps can enable an automatic quantification of brain tumor enhancement.

6.2 Methods

6.2.1 Dataset and preprocessing

UPenn-GBM dataset previously described in Chapter 1 was employed in this work. From all the image modalities, only the T1w and post-T1w were used. We selected a total of 220 patients (*UPenn-GBM-B*) with stable acquisition parameters. These 220 patients were randomly divided between training, early-stopping validation, and testing with an approximate proportion of 60%, 20%, and 20%, respectively.

6.2.2 Proposed approach

An overview of the training and testing approaches is shown in Figure 6.1. The CNN described in Chapter 4 was modified in order to compute the T1, T2, and PD parametric maps from only the T1w. Thus, the resulting CNN is composed of one encoder and three decoders, one per parametric map. Transfer learning was applied by initializing the network's weights with a pre-trained network with synthetic data following an approach similar to the one proposed in Chapter 3.

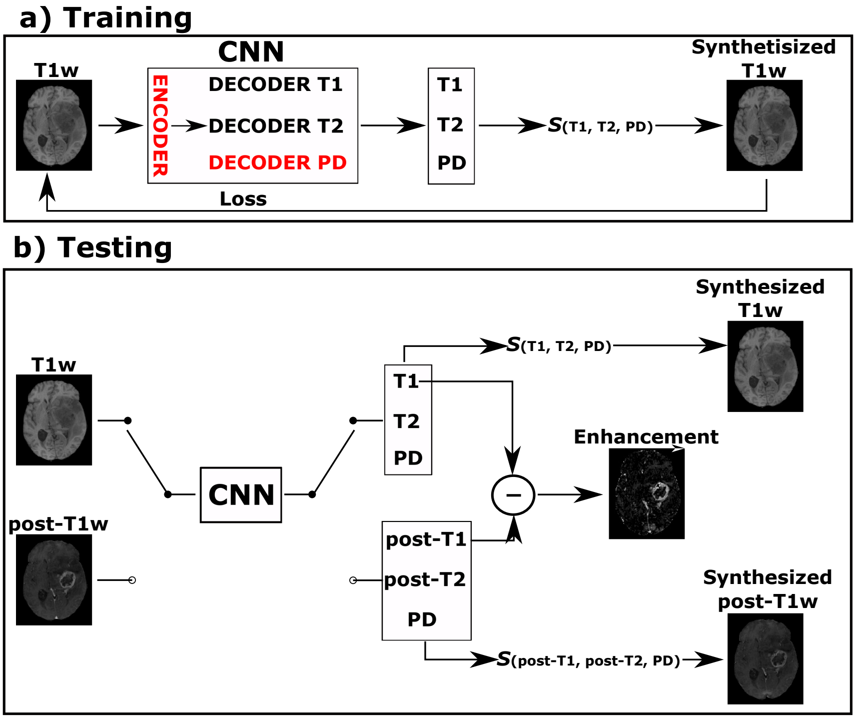


Figure 6.1: Overview of the proposed approach. a) Pipeline for training the CNN composed of an encoder and three decoders, one per parametric map. Transfer learning was applied by initializing the network's weights with a pre-trained network. Red parts were not re-trained. b) Pipeline for testing the trained CNN with the T1w and post-T1w inputs for the computation of pre- and post-contrast parametric maps, respectively. The difference between the T1 and post-contrast T1 (post-T1) maps corresponds to the enhancement quantification.

A self-supervised learning approach as proposed in Chapter 4 was employed to train the network only from the T1w images (i.e., in the weighted image domain) without the need of reference parametric maps. To this end, the T1w images were synthesized from the computed parametric maps using the theoretical pulse sequence expression of Eq. 2.10. The training was performed with a batch size of 32 and Adam optimizer (Jais et al., 2019) with learning rate of 1e-4. Early-stopping validation was employed to avoid overfitting.

For testing, the T1w and post-T1w were separately input to the trained network. Under the assumption that the only difference between both images is the GBCA intake (i.e., both images are acquired in the same scanner and with same acquisition parameters), the resulting maps should correspond to the pre- and post-contrast parametric maps, respectively. Thus, we can quantify the enhancement by means of the normalized difference between the pre-contrast T1 map and the post-T1 map. Specifically, the enhancement map could be computed as: $100 * \frac{|\mathrm{T1}-\mathrm{postT1}|}{\mathrm{postT1}} [\%]$.

Table 6.1: Quality metrics used to evaluate the performance of the T1w and post-T1w synthesis from the computed pre- and post-contrast parametric maps, respectively. Mean and SD values were computed across patients. Thus, the values are reported as mean ± SD. The metrics were calculated between both synthesized and acquired weighted images. The background pixels were not considered for the computation.

	MSE	**SSIM**	**PSNR**
T1w	0.0102 ± 0.0028	0.9730 ± 0.0035	27.9637 ± 1.8872
post-T1w	0.0261 ± 0.0093	0.9407 ± 0.0108	32.6785 ± 1.1742

6.3 Experimental work

First, the performance evaluation of the proposed approach was carried out by comparing the synthesized and acquired weighted images. The quality synthesis metrics employed were the MSE, SSIM, and PSNR computed between the synthesized and acquired T1w and post-T1w images. In addition, the pre-contrast and post-contrast T1 and T2 maps were visualized, and the enhancement maps were compared with the enhancement segmentation included in the *UPenn-GBM* dataset.

6.4 Results and discussion

Figure 6.2 shows representative slices of the synthesized and their corresponding acquired weighted images (both T1w and post-T1w) for different test patients. Table 6.1 shows the mean values along test patients of the quality metrics (i.e., MSE, SSIM, and PSNR) computed between the synthesized and the acquired weighted images for both image modalities. Figures 6.3 and 6.4 show for different test patients both the pre-contrast and post-contrast T1 and T2 maps, respectively. The corresponding enhancement maps are also shown in Figure 6.3. For the sake of visibility the intensity of the enhancement maps is cropped between 10% and 150%.

These results show the utility of parametric maps computed only from a routine sequence for the quantification of clinical relevant data. The synthesized weighted images can be visually compared with their actually acquired counterparts and visual resemblance is noticeable. The quality metrics also prove the agreement between both. Note that SSIM and PSNR present high values, whereas the values of the MSE are low.

Interestingly, within the enhanced region we can perceive the reduction of the T1 values in the post-contrast parametric maps compared to the corresponding pre-contrast maps. The enhanced region in the enhancement map considerably matches such a region in the segmentations computed from the weighted images. The shortening in post-T2 values can also be noticed in the enhancement region but to a much lower extent compared to the post-T1 values. Last, it is also worth mentioning the robustness of the approach for patients with and without tumor enhancement. See in Figure 6.3.c) that Patient 3 does not present enhancement.

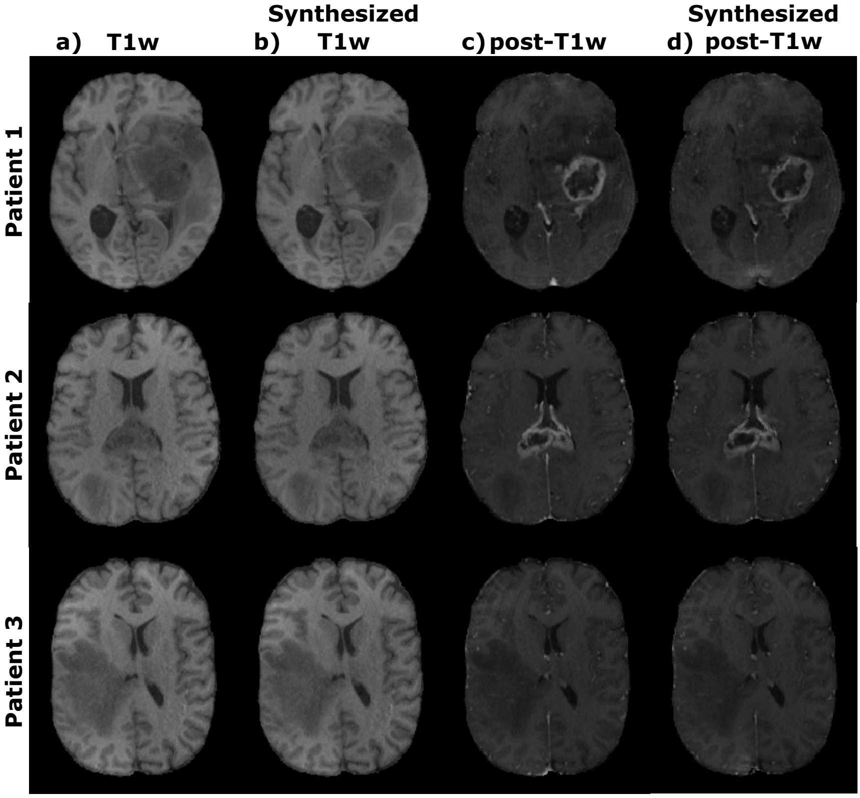


Figure 6.2: A representative axial slice of the T1w and post-T1w images synthesized from the computed parametric maps for three different test patients. a) Acquired T1w. b) Synthesized T1w. c) Acquired post-T1w. d) Synthesized post-T1w. Note that the T1w images are synthesized from the computed pre-contrast parametric maps, whereas the post-T1w images are synthesized from the computed post-contrast parametric maps.

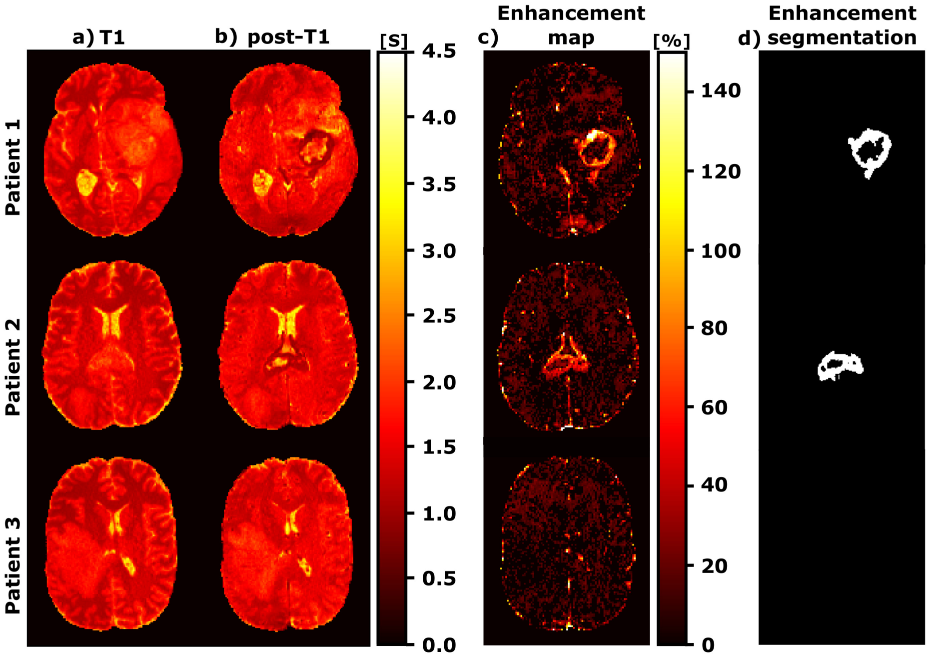


Figure 6.3: A representative axial slice of the T1, post-T1, and enhancement maps computed by the self-supervised CNN for three different test patients. a) T1 maps. b) post-T1 maps. c) Enhancement maps. d) Enhancement segmentation provided in the *UPenn-GBM* database. The T1 and post-T1 values are measured in seconds [s], whereas the enhancement maps values are measured in [%]. Note that Patient 1 and 2 present enhancement in the tumor but Patient 3 does not present such condition.

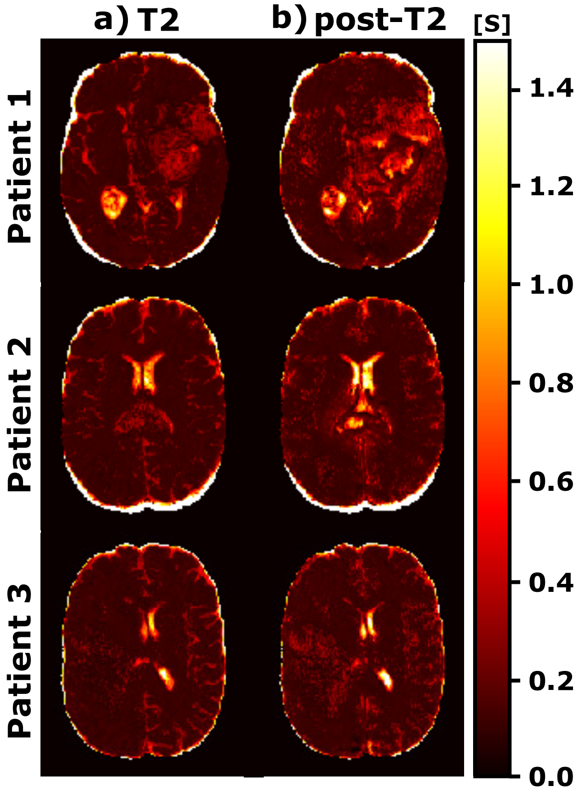


Figure 6.4: A representative axial slice of the T2 and post-contrast T2 (post-T2) maps computed by the self-supervised CNN for three different test patients. a) T2 maps. b) post-T2 maps. The T2 and post-T2 values are measured in seconds [s]. Similarly to the post-T1 map, the shortening in post-T2 values can be noticed in the enhancement region but to a lower extent.

The approach has some limitations; further validation of the computed parametric maps should be performed by comparing these maps with those obtained with other well-accepted yet lengthy relaxometry sequences. In addition, the value of the enhancement maps in clinical-decisions automatic pipelines should be studied.

6.5 Conclusions

In conclusion, in this chapter we proposed an approach for the enhancement quantification in glioblastoma by means of pre- and post-contrast parametric maps computed with DL from only a T1w and post-T1w obtained with routine fast sequences. Results suggest the potential of this approach for automatic quantitative assessment replacing the standard visual inspection of weighted images.

7

Post-contrast multiprametric mapping from only pre-contrast conventional weighted images

Contents

The work in this chapter has been partially published in:

- **Moya-Sáez, E.**, de Luis-Garcia, R., and Alberola-López, C. Toward deep learning replacement of gadolinium in neuro-oncology: A review of contrast-enhanced Synthetic MRI. *Frontiers in Neuroimaging.* 2023; 2, 1055463.
- **Moya-Sáez, E.**, de Luis-Garcia, R., Alberola-López, C., and Hernández-Tamames, J.A. Deep learning-based post-contrast imaging free of exogenous contrast agents. *3rd Annual Meeting of the ISMRM Iberian Chapter*, Valladolid, Spain, July 2023; P1.13.
- **Moya-Sáez, E.**, Nunez-Gónzalez, L., de Luis-Garcia, R., Alberola-López, C., and Hernández-Tamames, J.A. Post-contrast multi-parametric mapping from

only pre-contrast conventional weighted images. *2024 ISMRM & ISMRT Annual Meeting & Exhibition*, Singapore, May 2024; 2809.

7.1 Purpose

The acquisition of a weighted image after the injection a GBCA is common for brain tumor diagnosis and monitoring, and, in general, considered safe. However, up to 2.4% of injected patients suffer from mild adverse reactions and a lower rate for severe complication (Forghani, 2016). Also, the possible deposition of GBCAs in tissues has recently raised safety concerns (Gulani et al., 2017). In addition, the usage of GBCAs results in patient discomfort, additional cost and prolonged scan time. Thus, making GBCAs unnecessary would be highly advantageous. With this goal in mind, parametric maps could play a crucial role due to their known ability to detect subtle tumor changes (Hattingen et al., 2017; Nunez-Gonzalez et al., 2022).

In this chapter, we propose a DL approach for pre- and post-contrast parametric mapping and the synthesis of post-T1w images. The DL method is based on a cascade of two CNNs and it is trained *via* both supervised and self-supervised learning approaches. Only two pre-contrast conventional weighted images – a T1w and a T2w – acquired with routine sequences are used as input of the cascade of CNNs. Thus, this work paves the way towards the replacement of GBCAs for T1w-enhancement assessment.

7.2 Methods

7.2.1 Dataset and preprocesing

GLIOMA dataset (previusly described in Chapter 1) was employed in this work. This dataset includes four MR structural weighted images (T1w, T2w, T2w-FLAIR, and post-T1w) and, also, MAGiC* (Warntjes et al., 2008), which is a fast multiparametric mapping technique for the computation of T1, T2 and PD maps. The preprocessing steps are the same as in Chapter 5.

7.2.2 Proposed approach: *Cascade CNNs*

The pipeline we propose (see Figure 7.1) consists of a cascade of two CNNs, namely, an extraction CNN and a prediction CNN. The first CNN extracts the quantitative T1, T2, and PD maps whose information is embedded in the T1w and T2w input images. The second CNN takes these T1 and T2 maps and predicts the corresponding post-contrast (i.e., after GBCA administration) T1 and T2 maps. This prediction of the GBCA-related information is performed using both the tissue information derived from the pre-contrast maps and the extra knowledge extracted from the post-T1w and T2w during the self-supervised training. Training with T2w images, which are virtually invariant to contrast administration (Hattingen et al., 2017; Lescher et al., 2015), ensures that the post-contrast maps are numerically and physically coherent. Both CNNs share the same architecture as described in Chapter 3.

---

*Commercial implementation of the QRAPMASTER technique for quantitative imaging.

Training is two-fold; In step 1), the extraction CNN is trained via supervised learning with MAGiC parametric maps as reference. The loss function is the $l1$-norm defined as:

$$L = ||(\mathrm{T1}_c - \mathrm{T1}_{\mathrm{GT}})||_{\ell 1} + ||(\mathrm{T2}_c - \mathrm{T2}_{\mathrm{GT}})||_{\ell 1} + ||(\mathrm{PD}_c - \mathrm{PD}_{\mathrm{GT}})||_{\ell 1}, \tag{7.1}$$

where the subscripts c and GT refer to the DL-computed and the ground-truth (i.e., MAGiC) maps, respectively. Note that this equation corresponds with the loss function used to train the DL method of Chapter 5 (see Eq. 5.2.3).

In step 2), the prediction CNN is trained by means of a self-supervised approach as in Chapter 4, while keeping the weights of the extraction network fixed. For the self-supervised loss function we employ the theoretical pulse sequences expressions described in Eq. 2.10 and 2.11.

The loss function employed in step 2 is intended not only to achieve visual resemblance between the synthesized and the acquired post-T1w images but also to capture local characteristics within the tumor. Both the post-T1w and T2w are used as references. Consequently, the loss function is composed of three terms:

$$\begin{aligned} L_{\mathrm{step2}} = {} & ||(\text{post-T1w}_{\mathrm{syn}} - \text{post-T1w}_{\mathrm{acq}})||_{\ell 1} + ||(\mathrm{T2w}_{\mathrm{syn}} - \mathrm{T2w}_{\mathrm{acq}})||_{\ell 1} \\ & + \lambda ||M \odot (\Delta\mathrm{T1w}_{\mathrm{syn}} - \Delta\mathrm{T1w}_{\mathrm{acq}})||_{\ell 1} \end{aligned} \tag{7.2}$$

where $\text{post-T1w}_{\mathrm{syn}}$ and $\text{post-T1w}_{\mathrm{acq}}$ are the synthesized and acquired post-T1w, respectively. Similar distinction is applied to $\mathrm{T2w}_{\mathrm{syn}}$ and $\mathrm{T2w}_{\mathrm{acq}}$. M is a binary mask with positive values in the T1e regions, $\odot$ refers to Hadamard product, $\Delta\mathrm{T1w}_{\mathrm{syn}}$ and $\Delta\mathrm{T1w}_{\mathrm{acq}}$ are the GBCA intakes (i.e., the difference between the post-T1w — synthesized and acquired, respectively — and the acquired T1w). Finally, λ is a trade-off parameter to balance the contribution of the local loss. We stress that the post-T1w and the mask are only needed for training but not for testing.

Cross-validation was carried out *via* leave-one-out. For each data splitting, one patient is used for testing and the remaining patients are randomly split between training (11 patients) and early-stopping validation (2 patients). Both networks share the same patient splitting with the exception of the patient without the T2w-FLAIR, which was additionally included in the training set for the extraction network but not for the prediction one. Thus, no overly optimistic splitting was performed.

Transfer learning techniques were employed in both networks. In the extraction CNN layers weights were initialized by training with a purely synthetic dataset as in Chapter 3. In the prediction network weights were initialized by pretraining with a larger public dataset (Bakas et al., 2022) of glioblastoma patients with no resection. Only the decoders were re-trained in the prediction CNN. Additionally, data augmentation with horizontal flips was included for the training of the prediction CNN. Both networks were trained with Adam optimizer (Jais et al., 2019) with learning rate of $1e^{-4}$ and early-stopping. Parameters empirically set were the batchsize = 4 (in both networks) and λ=100.

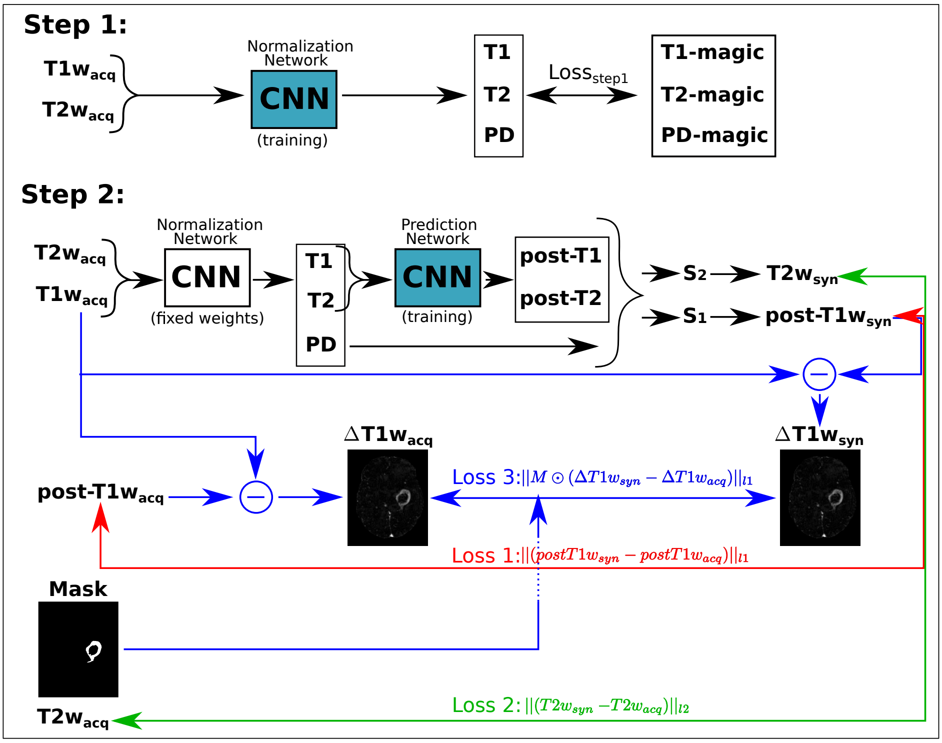


Figure 7.1: Overview of the proposed approach. The cascade of CNNs is composed of an extraction CNN and a prediction CNN. In each step the colored network is trained. In step 2 the color of the arrows represents the computations related to each loss term. Loss term 1 (red) focuses on achieving resemblance between the acquired and synthesized post-T1w, loss term 2 (green) ensures physical coherent values on post-contrast parametric maps, and loss 3 (blue) tries to capture local characteristics within the tumor.

7.3 Experimental work

The impact that the synthesized post-T1w images could have in clinical practice should be evaluated with quantitative algorithms for improving clinical decision-making by predicting tumor enhancement. HD-GLIO (Kickingereder et al., 2019) is a DL segmentation tool that automatically segments voxels with T1e and voxels with T2h — jointly referred to as ABN —. This tool is fed with four weighted images (i.e., T1w, T2w and T2w-FLAIR, as well as post-T1w). For our evaluation, we used the three pre-contrast acquired weighted images, as well as the synthesized post-T1w to segment T1e and T2h regions with HD-GLIO. These segmentations are compared with the ground-truth segmentations obtained with HD-GLIO when fed with the four acquired weighted images. In addition, the results obtained were compared with another recently proposed approach for T1w-enhancement prediction (Nunez-Gonzalez et al., 2022). This approach carries out a voxel-wise classification using pre-contrast parametric maps obtained with MAGiC (hereinafter referred to as *Voxel-wise MAGiC*). In this work, *Voxel-wise MAGiC* was computed with the best metric and threshold reported in Nunez-Gonzalez et al. (2022) for classification problems C-I and C-II defined in Chapter 5 and Appendix D. Note that the output of these two classifications are equivalent to the T1e and T2h segmentations.

Thus, our evaluation consists in comparing closeness of both our method (*Cascade CNNs*) and *Voxel-wise MAGiC* to the ground-truth. The accuracy of both segmentations with respect to the ground-truth is separately measured voxel-wise and lesion-wise as we now describe:

Voxel-wise evaluation. We voxel-wise measured the sensitivity and specificity of classification problems C-I and C-II for both *Voxel-wise MAGiC* and our method *Cascade CNNs*.

Lesion-wise evaluation. Images that show the presence of enhancement in the surroundings of the tumor area, even though the enhancement they show does not cover the whole tumor, could be valuable for decision-making. To reflect this idea, we have designed a lesion-wise performance measurement consisting in: 1) the computation of clusters of voxels whose size are larger than 10% of the total volume of T1e in the ground-truth mask; and 2) the computation of the sensitivity considering a true positive when any of these clusters are, totally or partially, included within the ground-truth T1e segmentation. This metric is computed for both *Voxel-wise MAGiC* and our method *Cascade CNNs*.

7.4 Results and discussion

Figure 7.2 shows the pre- and post-contrast parametric maps for different test patients. Interestingly, as in previous Chapter 6, a shortening in the T1 values can be noticed within the enhancement region, which is in agreement with the literature (Hattingen et al., 2017; Lescher et al., 2015; Warntjes et al., 2018). This behaviour also occurs in the T2 values but to a much lower extent. The PD map is

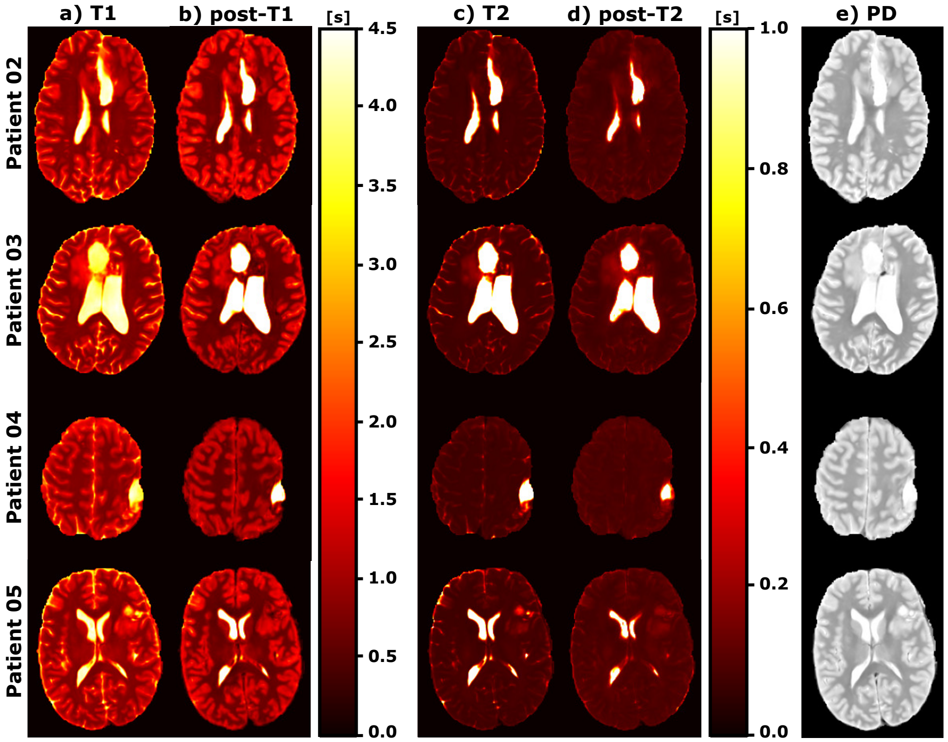


Figure 7.2: A representative axial slice of the pre-contrast and post-contrast computed parametric maps on different test patients of *GLIOMA*. a) Pre-contrast T1 maps. b) Post-contrast T1 maps. c) Pre-contrast T2 maps. d) Post-contrast T2 maps. e) PD maps. The pre- and post-contrast T1 and T2 values are measured in seconds [s].

not affected by the GBCA injection. Note that the computed T1, T2 and PD maps correspond to the output of the extraction CNN (step 1 in Figure 7.1).

Figure 7.3 shows both the synthesized and actually acquired post-T1w images for different test patients of *GLIOMA* dataset. The corresponding ground-truth and synthetic HD-GLIO segmentations are also shown overlaid on the acquired post-T1w and synthesized post-T1w, respectively. Note that the post-T1w images are synthesized from the computed post-contrast parametric maps shown in Figure 7.2 using the theoretical equation that describes MR image intensity. The synthesized post-T1w images can be visually compared with their actually acquired counterparts and visual resemblance is noticeable.

Results of voxel-wise and lesion-wise evaluations are shown in Tables 7.1 and 7.2, respectively. Table 7.1 shows a superiority of *Voxel-wise MAGiC* in terms of sensitivity, although lower specificity values are reported with this method. As for our method, we have observed that the enhanced areas in the synthesized images do not show a by-point coincidence with ground-truth despite these images indeed show the presence of enhancement. This fact can be appraised in the figures of Table 7.2 where *Cascade CNNs* achieves comparable performance with *Voxel-wise*

Table 7.1: Sensitivity and Specificity of the voxel-wise predictions for both classification problems (i.e., C-I and C-II) and both methods — *Voxel-wise MAGiC* and *Cascade CNNs* —. Median across patients and IQR are reported.

	(C-I) ABN vs. nWM		**(C-II)** T1e vs. non T1e	
	Voxel-wise MAGIC	***Cascade CNNs***	***Voxel-wise MAGIC***	***Cascade CNNs***
Sensitivity	89.35% (IQR: 8.93%)	77.92% (IQR: 21.59%)	87.29% (IQR: 16.62%)	60.20% (IQR: 26.20%)
Specificity	93.61% (IQR: 2.14%)	99.95% (IQR: 0.13%)	94.95% (IQR: 2.88%)	99.40% (IQR: 0.66%)

Table 7.2: Sensitivity of the lesion-wise enhancement prediction for each patient and both methods. Note that only the patients with enhancing lesions are considered. Bold text represents the superiority of the *Cascade CNNs* method, whereas inferiority is represented with red text.

Patient ID	***Voxel-wise MAGiC***	***Cascade CNNs***
00	33%	**100%**
01	40%	**50%**
02	100%	100%
03	100%	50%
04	75%	50%
05	33%	**100%**
07	100%	100%
08	100%	100%
09	100%	100%
12	100%	100%
13	50%	0%

MAGiC. In addition, *Cascade CNNs* provides the post-contrast parametric maps as a byproduct.

These preliminary results show how using only pre-contrast weighted images (i.e., acquired without GBCAs) is feasible not only to extract pre-contrast parametric maps but also to predict post-contrast maps. These maps may add value towards automatic enhancement quantification (Hattingen et al., 2017; Lescher et al., 2015). Additionally, the predicted post-contrast maps have the capability of synthesizing post-contrast weighted images with visual resemblance to the acquired counterparts as shown in Figure 7.3. Despite the differences visible in Figure 7.3 between the ground-truth and the synthetic T1w-enhancement segmentations, the clinical value of both could potentially be comparable in terms of their capacity to detect lesions.

Future lines of research include the comparison of the computed post-contrast parametric maps with those obtained with other well-accepted techniques as well as the validation of the approach with a larger cohort.

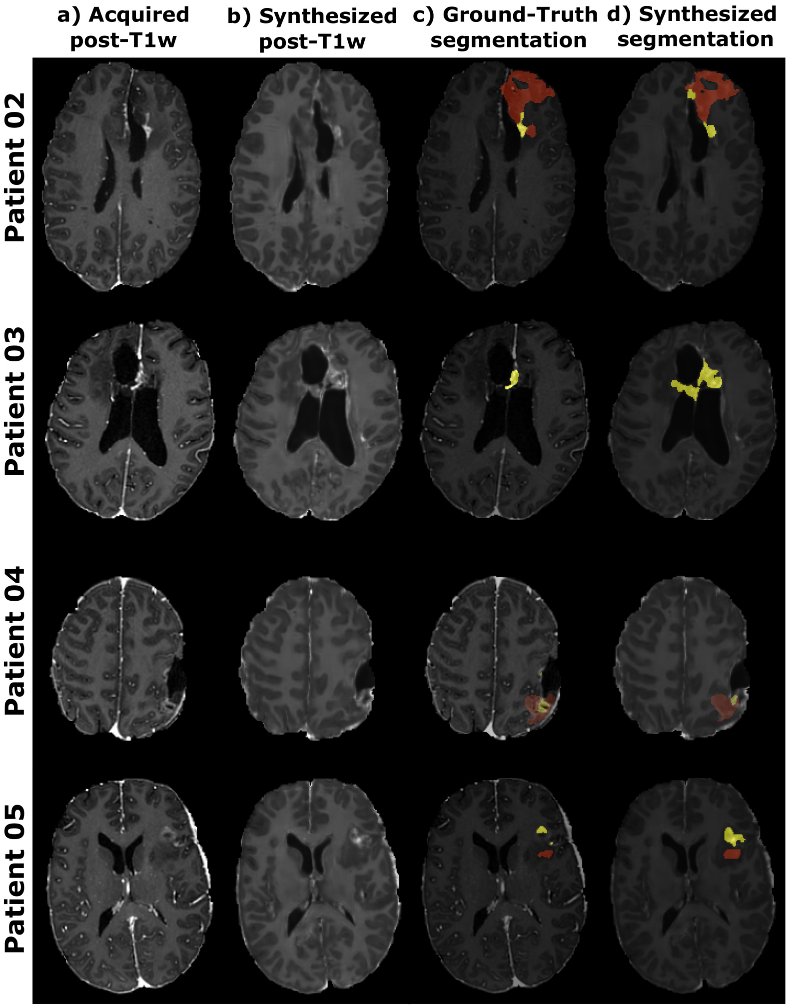


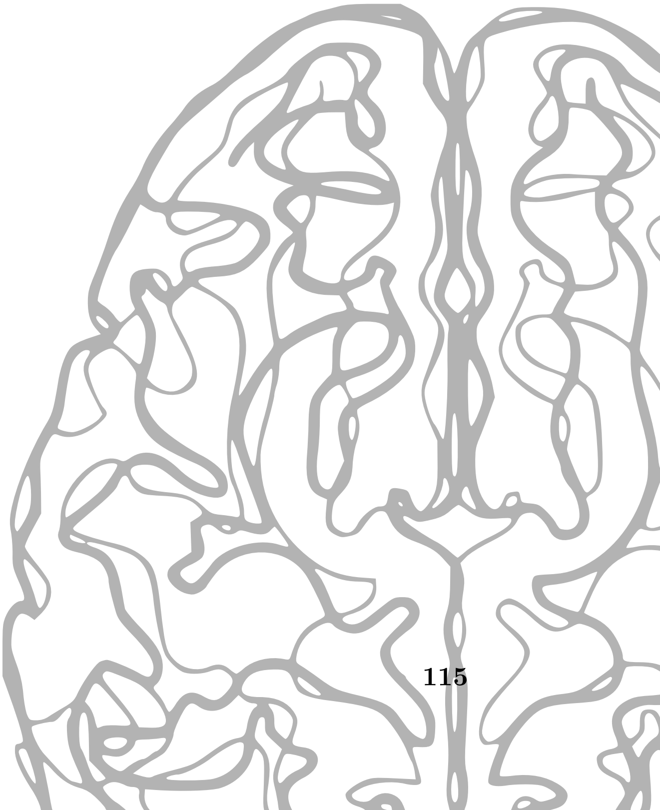
Figure 7.3: A representative axial slice of the synthesized post-T1w images for different test patients. a) Acquired post-T1w. b) Synthesized post-T1w. c) Ground-truth segmentation mask of HD-GLIO. d) Corresponding segmentation mask of HD-GLIO using the synthesized post-T1w instead of the acquired one. Yellow label: T1e, red label: T2h.

7.5 Conclusions

In conclusion, in this chapter a novel approach for pre- and post-contrast parametric mapping and the synthesis of post-T1w images is proposed. The computation is based on a cascade of CNNs and only needs a pair of pre-contrast conventional weighted images as input. Our results suggest the potential of this approach for replacing GBCAs for T1w-enhancement assessment.

Part V

Conclusions



8

Conclusions, future work and contributions

Contents

8.1 Conclusions

In this Thesis we have described several works which aim at improving the diagnosis of brain tumors following a Synthetic MRI paradigm. As previously described, this paradigm is based on the computation of T1, T2, and PD parametric maps and, also, the synthesis of several weighted images out of the computed parametric maps. Parametric maps, as described in the literature and as our own results suggest, add value to the diagnosis due to their quantitative nature. Hence, they are key ingredients for distinguishing different post-treatment conditions, understanding biological changes within the tumor and minimizing interpretation errors of treatment effects (Deoni, 2010; Hattingen et al., 2017; Pirkl et al., 2021). However, since radiologists are used to perform the diagnosis based on weighted images (Ellingson et al., 2015a), the synthesis of this kind of images is also of utmost importance and might lead to MR protocol shortening and an increase of the patient well-being.

As we have shown, the computation of the parametric maps can be leveraged by the usage of DL methods. Hence, parametric maps can be computed from conventional weighted images instead of lengthy relaxometry or scarcely available specific sequences. In order to overcome the lack of large, public datasets with paired weighted images and parametric maps to train the DL methods, a novel strategy based on training with synthetic images was proposed. To improve the computation accuracy, the network could be fine-tuned with real data, but, in any case, the synthetic training has proven to be crucial for obtaining realistic map computations with small training datasets.

In addition, a self-supervised learning approach has been proposed by introducing the theoretical pulse sequences equations in the loss function. Thus, DL methods can be trained with weighted images instead of parametric maps. Note here that weighted images are much easier to obtain than parametric maps.

Our results demonstrate the utility of including synthesized weighted images for feeding a radiomics system for predicting the survival of glioblastoma patients. Specifically, the radiomics system was tested replacing one acquired input image with its synthesized counterpart. We also compared performance with a radiomics system trained from scratch ignoring this additional channel. The performance achieved with the synthesized images in the radiomic system is not only close to the performance achieved using all the acquired images, but also substantially outperformed a model trained without that weighting. These findings indicate that the usage of synthesized images could facilitate the widespread usage of multicontrast-demanding quantitative applications, such as radiomics, by retrospectively completing databases with missing modalities and/or replacing artifacted images. Further, synthesized images have the potential to speed up acquisition protocols by replacing some acquired weighted images with their synthesized counterparts.

Promising results were also observed in the detection of tumor T1w-enhancement. In this line of research, parametric maps could be a potential solution for both enabling automatic quantification of T1w-enhancement as opposed to standard visual inspection of weighted images, and replacing the usage of GBCAs. Regarding the latter, parametric maps computed with DL from conventional weighted images show similar performance to discriminate normal/abnormal and T1w-enhanced/non-enhanced tissue than parametric maps obtained with MAGiC or relaxometry sequences. These results could be replicated by testing the DL-based method with additional multi-center, multi-vendor datasets unseen during the training process. The computation of parametric maps from conventional weighted images, as proposed in this Thesis, would promote the adoption of this clinical application to predict T1w-enhancement without GBCAs and, in a broader sense, the usage of parametric maps in the clinical practice. Moreover, this approach proves valuable for retrospective analysis, enabling the computation of parametric maps for pre-existing and/or multicenter datasets, thereby providing means for boosting longitudinal or population studies.

Finally, we have shown preliminary results on how using only pre-contrast weighted images it is feasible not only to extract pre-contrast parametric maps but also to predict post-contrast maps. Once again, these maps may add value towards automatic enhancement quantification. Additionally, the predicted post-contrast maps have the capability of synthesizing post-contrast weighted images with visual resemblance to the acquired counterparts. Despite the differences between the ground-truth and the synthetic T1w-enhancement segmentations, the clinical value of both could potentially be comparable in terms of their capacity to detect lesions.

8.2 Future work

8.2.1 Technical considerations

Throughout the development of this Thesis, we have encountered certain limitations that open up future lines of research for extending and improving the proposed approaches.

The theoretical pulse sequences equations used to synthesize the weighted images did not consider all the effects that occur in practice. For example, the T2*w image in Chapter 3 is synthesized from the T2 map instead of the T2* map. This aspect could have an impact on the proposed approach and explain the worse metrics for this weighted image compared to the metrics of the T2w image. Also, in that chapter the T2w-FLAIR presents worse metrics than the other modalities, although the work in Chapter 4 focuses on improving the synthesis quality of this weighted image. Overall, a more realistic synthesis of weighted images with a detailed Bloch simulation (Cao et al., 2014; Castillo-Passi et al., 2023; Stöcker et al., 2010) and other maps (e.g. T2*, B0, and B1 maps) could also enhance the synthesis quality of any weighted image modality. Further, as we previously suggested in Chapter 3, the simulation of motion in the synthetic dataset could enhance the network robustness against such artifacts. A complementary approach to improve the computation and synthesis performance could be the selection of the optimal pulse sequences and/or sequence parameters to feed the DL method.

The self-supervised approaches are dependent on the acquisition parameters of the weighted images. This has the advantage of making the process specific to this parameter setting, so higher synthesis quality could be expected. The downside is the inherent limitation to this particular setting. Nevertheless, the self-supervised method can be easily extended to accommodate more parameter values for which acquisitions are available. In some cases as in Chapter 5, the T1w and T2w input images have varying acquisition parameters across datasets due to different institutional protocols. Although this study has indeed the value of showing comparable results when using other datasets for testing, these input parameter variations can potentially introduce errors in the computation of the parametric maps. To mitigate the impact of variation in acquisition parameters, future work aims to incorporate them as additional inputs to the network. This approach would enable training with a more extensive and diverse dataset, potentially enhancing the model generalization capabilities.

In general, scan time limitations lead to a mismatch in spatial resolution between the weighted images and the parametric maps. Consequently, the co-registration of all modalities is needed, which might induce misregistration artifacts and partial volume effects. Specifically, misregistration of the input weighted images might be one of the main sources of errors of the proposed DL methods. However, in production mode the input images should be ideally acquired with the same spatial resolution minimizing these issues.

The T1w-enhancement prediction in Chapter 5 could be performed directly from the pre-contrast weighted images instead of the computed parametric maps. However, we choose to use the maps as an intermediate step to ground our method on previous studies, which have shown that quantitative parametric maps support the rationale of T1w-enhancement prediction when replacing GBCAs, so that this DL finding can be clinically explained. Following the same criteria, the survival prediction of glioblastoma patients proposed in Chapter 4 could be performed directly on the parametric maps. The usage of a radiomics system fed with weighted images was a design choice, but, indeed, trying to perform this prediction from the parametric maps is highly interesting and could be a possible future line of research. Further research to analyze the potential of parametric maps as follow-up biomarker of the impaired BBB, as presumable show our results of Chapter 5, is also strongly recommended.

Experiments on larger and multi-institutional cohorts including parametric maps would be advisable to further support our conclusions. Specifically, in Chapters 6 and 7 that focused on the computation of post-contrast parametric maps, the comparison of the DL-computed post-contrast maps with those obtained with other well-accepted techniques, such as relaxometry or multiparametric mapping, is highly desirable to confirm the potential of our preliminary findings.

Additional future work includes improvement and further tuning in the implemented CNN and/or implementing other advanced NN architectures, such as those introduced in Subsection 2.5.5 (i.e., GANs and ViTs architectures). Particularly, the usage of GAN architectures may be studied due to the recent works that achieve impressive results in medical image translation capturing high-frequency texture information (Armanious et al., 2020; Dar et al., 2019). Also ViTs are novel architectures that, by incorporating attention mechanisms, could play an important role in Synthetic MRI (Dalmaz et al., 2022).

A direction for future research could also be the extension of the proposed Synthetic MRI approaches to other slice orientations, such as sagittal or coronal, and/or other tissues, such as the heart or the liver. To this end, for example, the extended Cardiac-Torso (XCAT) phantoms (Segars et al., 2010) could be employed. Furthermore, the extension of the approach to predict T1w-enhancement without GBCAs for diagnosis is also planned.

8.2.2 Novel potential application

Low-field MRI: Low-field MRI systems, operating at field strengths typically lower than 1.5T (Arnold et al., 2023), offer several advantages such as lower cost and power consumption, increased patient accessibility, and fewer safety concerns. Still, these systems have to face several limitations derived from their lower signal availability (reduced SNR, lower resolution, increased acquisition time, etc.) (Arnold et al., 2023). Thus, the lower magnetic field strength leads to decreased sensitivity, resulting in poorer image quality and reduced conspicuity of lesions or abnormalities. Undersampling techniques followed by DL reconstruction methods have reduced acquisition time and improved image quality making low-field MRI sufficient for some clinical applications (Hori et al., 2021; Iturri-Clavero et al., 2016). Synthetic MRI methods could also play a key role by the synthesis of different high-quality image modalities from a short protocol. However, it is important to note that the relaxivity differences compared to high-field MRI affect the values of T1 and T2 and, thus, these methods have to be applied carefully.

Additional limitations of low-field MRI are related to the usage of GBCAs. The relaxivity of GBCAs, which determines their efficacy in T1 shortening and enhancing contrast, diminishes at lower field strengths. This decrease in relaxivity reduces the enhancement achieved, making it challenging to visualize subtle pathologies or small lesions (Brekenfeld et al., 2001) and it could affect procedures where resection of enhancing tissue predicts patient outcome (Arnold et al., 2023; Hori et al., 2021). Significantly, it has been observed that enhancing contrast at low field strength with a double dose (0.2 mmol/kg) of intravenous GBCA administration yields comparable results to those achieved at 1.5T with the standard dose (0.1 mmol/kg) (Desai and Runge, 2003). However, as it was previously explained in this Thesis, increasing GBCAs dosages could raise safety risks. As a result, in low-field MRI there is a need for alternative imaging strategies for effective diagnostic evaluation of T1w-enhancement which could be overcome by Synthetic MRI methods without GBCAs similar to the methods proposed in Part IV of this Thesis.

Tumor micro-structure: Diffusion-weighted magnetic resonance imaging (dMRI) has emerged as a prevalent non-invasive method for assessing tissue micro-structure within tumors (Zhu et al., 2023). This analysis opens avenues in cancer diagnosis, tumor grading, and evaluating treatment efficacy (Fokkinga et al., 2023). The current Synthetic MRI approach for computation of T1, T2, and PD parametric maps can be strengthened by utilizing diffusion sequences within a joint modeling framework (Coelho et al., 2024; Wang et al., 2023). Additionally, this approach could also be modified to generate micro-structure maps instead of relaxometry parametric maps (Chiou et al., 2021; Yılmaz et al., 2024). Although the incorporation of these type of prior physical knowledge could be challenging, several works have proposed self-supervised approaches to estimate diffusion parameters (Ehrlich and Rivera, 2021; Våge, 2023; Vasylechko et al., 2022). After computing these diffusion parameters, the third step of the Synthetic MRI (see Section 2.4) could be modified so that diffusion sequences can be synthesized from existing maps to facilitate diagnosis, enabling adjustment of b-values and synthesis of specific sequences as needed (Hu et al., 2021).

8.3 Contributions

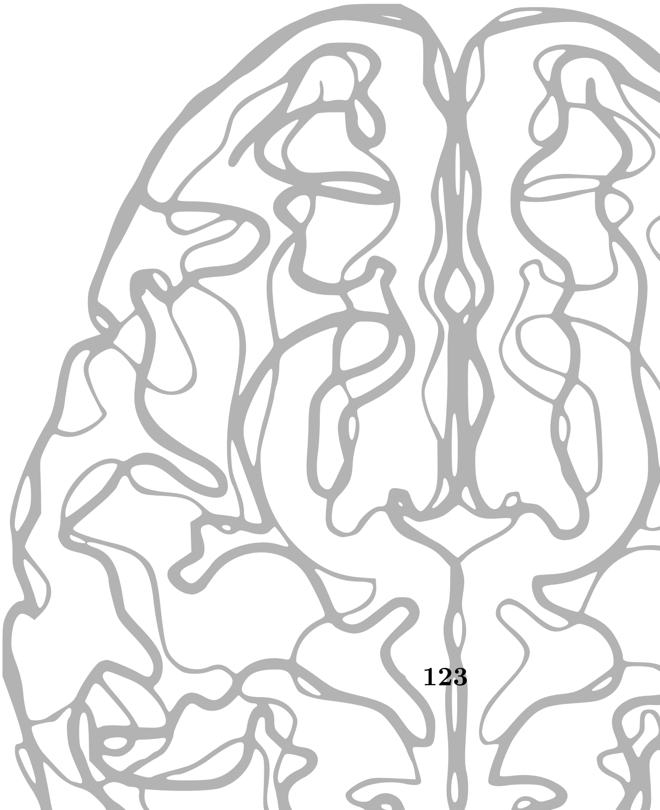
In this Thesis we have focused on the usage of a Synthetic MRI paradigm to enhance the diagnosis of different types of brain tumors. The computation of T1, T2, and PD parametric maps has been performed with DL from conventional weighted images acquired with routine protocols. After that, different types of weighted images have been successfully synthesized out of the parametric maps. Both the synthesized weighted images and the computed parametric maps have been employed in different application for improving tumor diagnosis; specifically, predicting both the expected survival of glioblastoma patients and the post-contrast T1w-enhanced tissues without the injection of GBCAs.

The main scientific contributions are listed next:

- Development of a novel DL approach for the computation of T1, T2, and PD parametric maps from only a pair of conventional weighted images (a T1w and a T2w). In addition, a new training strategy based on a synthetic dataset generated from the BrainWeb anatomical brain model is proposed. This way, the need of large datasets with quantitative parametric maps is bypassed.
- Synthesis of multiple realistic weighted images from these computed parametric maps both for modalities previously seen by the network and for other modalities not used in the learning stage. Quantitative and qualitative comparisons between the synthesized and the acquired weighted images are provided.
- A self-supervised approach for training DL methods for multiparametric mapping and synthesis of weighted images in the weighted images domain (i.e., without the need of reference parametric maps for training).
- Provision of empirical evidence that synthesized weighted images can be used for feeding a radiomics system for survival prediction of glioblastoma patients achieving similar performance as with using the acquired ones. Moreover, our method shows the advantages derived from the synthesis itself, such as the possibility of retrospectively completing databases with missing modalities and/or replacing artifacted images.
- Extension of the proposed self-supervised Synthetic MRI approach for the computation of both pre- and post-contrast parametric maps for automatic quantification of tumor enhancement.
- Confirmation that parametric maps computed with DL can be used as a surrogate of traditional relaxometry sequences or other commercial techniques, such as MAGiC, for predicting brain tumor T1w-enhancement without GBCAs
- First steps towards the avoidance of GBCAs in post-contrast T1w imaging acquisitions with, in our opinion, promising results.

Part VI

Appendices



A

MR acquisition details

This appendix includes details about the acquisition parameters for the image modalities included in the different datasets employed in this Thesis. Specifically, these parameters can be found in Table A.1 for *Multicontrast Brain-MRI*, Table A.2 for *Relaxometry Brain-MRI*, Table A.3 for *Multicontrast Glioblastoma*, Table A.4 for *Upenn-GBM*, and Table A.5 for *GLIOMA*.

Table A.1: Acquisition parameters for *Multicontrast Brain-MRI* dataset.

	Multicontrast Brain-MRI			
	T1w (IR-GRE)	**PDw/T2w** (ME-TSE)	**T2w-FLAIR** (IR-TSE)	**T2*w** (GRE)
TE (ms)	3	30/85	100	20
TR (ms)	6.44	4000	11000	746.99
TI (ms)	900	-	2800	-
α (°)	10	-	-	20
# Slices (Orientation)	170 (sagittal)	50 (axial)	27 (axial)	27 (axial)
Slice thickness (mm)	1.2	3	5	5
Voxel size (mm^2)	1.25 × 1.25	1.02 × 1.36	0.94 × 1.25	0.94 × 1.25
FOV (mm)	240 × 240	260 × 195	240 × 240	240 × 240
Scan time (min)	$\sim$ 4:00	2:30 - 4:30	2:30 - 4:30	2:30 - 4:00

Table A.2: Acquisition parameters for *Relaxometry Brain-MRI* dataset.

	Relaxometry Brain-MRI				
	T1 map (VFA GRE)	**T2 map** (ME-SE)	**PD map** (GRE)	**T1w** (MPRAGE)	**T2w** (TSE)
TE (ms)	2	17,46,75,104,133,162	2	3	85
TR (ms)	18	1000	50	6.44	4000
TI (ms)	-	-	-	900	-
α (°)	2, 3, 4, 5, 7, 9, 11, 14, 17, 19, 22	-	5	10	90
# Slices (Orientation)	150 (axial)	150 (axial)	150 (axial)	170 (sagittal)	50 (axial)
Slice thickness (mm)	1.5	1.5	1.5	1.2	3
Voxel size (mm^2)	1.50 × 1.50	1.50 × 1.50	1.50 × 1.50	1.25 × 1.25	1.02 × 1.36
FOV (mm)	240 × 240	240 × 240	240 × 240	240 × 240	260 × 195
Scan time (min)	$\sim$ 17:00	$\sim$ 18:00	4:00 - 4:30	$\sim$ 4:00	2:30 - 4:30

Table A.3: Acquisition parameters for *Multicontrast Glioblastoma* dataset. UN stands for unknown.

	Multicontrast Glioblastoma			
	T1w (FSPGR)	**T2w** (TSE)	**T2w-FLAIR** (IR-TSE)	**post-T1w** (FSPGR)
TE (ms)	1.83	122	142	2.18
TR (ms)	5.98	4162	9350	6.85
TI (ms)	-	-	2200	-
α (°)	12	90	-	12
# Slices	216	32	32	216
Slice thickness(mm)	1.6	4	4	1.6
Voxel size (mm^2)	0.94 × 0.94	0.78 × 0.87	0.7 × 1.0	0.94 × 0.94
FOV (mm)	240 × 240	240 × 240	240 × 240	240 × 240
Scan time (min)	UN	UN	UN	UN

Table A.4: Acquisition parameters for *UPenn-GBM* dataset. UN stands for unknown.

	UPenn-GBM			
	T1w (MPRAGE)	**T2w** (SE)	**T2w-FLAIR** (IR-TSE)	**post-T1w** (MPRAGE)
TE (ms)	3.1	458	140	3.1
TR (ms)	1760	3200	9420	1760
TI (ms)	950	-	2500	950
α (°)	15	120	170	15
# Slices	155	155	155	155
Slice thickness (mm)	1.0	0.9	3.0	1.0
Voxel size (mm^2)	0.98 × 0.98	0.90 × 0.90	0.94 × 0.94	0.98 × 0.98
FOV (mm)	240 × 240	240 × 240	240 × 240	240 × 240
Scan time (min)	UN	UN	UN	UN

Table A.5: Acquisition parameters for *GLIOMA* dataset.

	GLIOMA				
	T1w	**T2w**	**T2w-FLAIR**	**post-T1w**	**MAGiC**
	(IR-GRE)	(TSE)	(IR-TSE)	(IR-GRE)	(MDME)
TE (ms)	3.3	97	89	.3	6114
TR (ms)	7.9	9837	5000	7.9	15.7
TI (ms)	450	-	1588	450	11
α (°)	12	90	90	12	90
# Slices **(Orientation)**	352 (axial)	49 (axial)	224 (sagittal)	352 (axial)	49 (axial)
Slice thickness (mm)	1.2	3.0	1.6	1.0	3.0
Voxel size (mm^2)	1.0 × 1.0	0.6 × 0.6	1.11 × 1.11	1.0 × 1.0	1.0 × 1.0
FOV (mm)	240 × 240	233 × 233	246 × 246	240 × 240	240 × 240
Scan time (min)	∼ 5:00	∼ 4:00	∼ 4:00	∼ 5:00	∼ 5:00

B

Inversion recovery-Gradient echo (IR-GRE) theoretical pulse sequence equation

In order to obtain the theoretical expression for this pulse sequence some assumptions need to be taken under consideration. We consider that the stady-state has been reached. M_{ss} is the transversal component of the magnetization before the inversion pulse at the time $n-1$. Additionally:

- $I^{n}_{ss}(0_{-})$: is the longitudinal magnetization component just before the inversion pulse of θ_{inv} radians in the n-th cycle with $(n \leq 0)$. $I^{n}_{ss}(0_{+})$ is the equivalent but just after the inversion pulse.

- $E^{n}_{ss}(0_{-})$: is the longitudinal magnetization component just before the excitation pulse of θ radians in the n-th cycle with $(n \leq 0)$. $E^{n}_{ss}(0_{+})$ is the equivalent but just after the excitation pulse.

- $\lambda = e^{-TI/T_1}$; $\beta = e^{-(T_R-TI)/T_1}$ and, consequently, $\lambda\beta = e^{-(T_R)/T_1}$.

Under these assumptions:

- Longitudinal component:

$$
\begin{aligned}
I_{ss}^{n-1}(0_+) &= M_{ss}\cos(\theta_{inv}) \\
E_{ss}^{n-1}(0_-) &= M_{ss}\cos(\theta_{inv})\lambda + M_z^0(1-\lambda) \\
E_{ss}^{n-1}(0_+) &= E_{ss}^{n-1}(0_-)\cos(\theta) \\
I_{ss}^{n}(0_-) &= E_{ss}^{n-1}(0_+)\beta + M_z^0(1-\beta) = E_{ss}^{n-1}(0_-)\cos(\theta)\beta + M_z^0(1-\beta) \\
I_{ss}^{n}(0_+) &= I_{ss}^{n}(0_-)cos(\theta_{inv}) = \big(E_{ss}^{n-1}(0_-)\cos(\theta)\beta + M_z^0(1-\beta)\big)\cos(\theta_{inv}) \\
E_{ss}^{n}(0_-) &= I_{ss}^{n}(0_+)\lambda + M_z^0(1-\lambda) \\
&= \big(E_{ss}^{n-1}(0_-)\cos(\theta)\beta + M_z^0(1-\beta)\big)\cos(\theta_{inv})\lambda + M_z^0(1-\lambda)
\end{aligned}
$$

  Considering that $E_{ss}^{n}(0_-) = E_{ss}^{n-1}(0_-)$, we obtain:

$$
\begin{aligned}
E_{ss}^{n}(0_-) &= \big(E_{ss}^{n}(0_-)\cos(\theta)\beta + M_z^0(1-\beta)\big)\cos(\theta_{inv})\lambda + M_z^0(1-\lambda) \\
E_{ss}^{n}(0_-) &= M_z^0\frac{1-\lambda+(1-\beta)\lambda\cos(\theta_{inv})}{1-\cos(\theta)\cos(\theta_{inv})\beta\lambda} \\
&= M_z^0\frac{1-e^{-TI/T_1}+(1-e^{-(T_R-TI)/T_1})e^{-TI/T_1}\cos(\theta_{inv})}{1-\cos(\theta)\cos(\theta_{inv})e^{-T_R/T_1}} \\
&= M_z^0\frac{1-(1-\cos(\theta_{inv})e^{-TI/T_1}-e^{-T_R/T_1}\cos(\theta_{inv})}{1-\cos(\theta)\cos(\theta_{inv})e^{-T_R/T_1}}
\end{aligned}
$$

- Transversal component:under the assumption that the residual transversal components are spoiled as a consequence of spoilers gradients. In that case, the transversal magnetization at the TE should correspond with:

$$
M_{xy}^{n}(T_E) = M_z^0\frac{1-(1-\cos(\theta_{inv})e^{-TI/T_1}-e^{-T_R/T_1}\cos(\theta_{inv})}{1-\cos(\theta)\cos(\theta_{inv})e^{-T_R/T_1}}\sin(\theta)e^{-TE/T_2^*} \tag{B.1}
$$

C

Radiomics for glioblastoma survival prediction

Contents

C.1 Datasets

Three different datasets of glioblastoma patients were used for training the radiomics system. Two of them are publicly available, namely, the *BraTS2020* (Multimodal Brain Tumor Segmentation) (Menze et al., 2014) 2020 Challenge dataset, and the datasets reachable through *TCIA* (The Cancer Image Archive) (Clark et al., 2013) — which, in turn, consist of three sources, namely, Ivy Glioblastoma Altas Project (Ivy-GAP), the Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme (CPTAC) and The Cancer Genome Atlas Glioblastoma Multiforme (TCGA)—. The other dataset (*Dataset22*), is a private dataset acquired in Hospital Universitario Río Hortega, Valladolid, Spain. These datasets where only used with the purpose of training the radiomic system, hence, they are not included in Section 1.4 of this Thesis dissertation.

For testing the radiomics system in combination with the synthesis method, we employed a private dataset (*Multicontrast Glioblastoma*), which was introduced in Section 1.4 of Chapter 1.

Details of the datasets can be found in Table C.1. From all the datasets, we only included those patients (199 patients in total) in which gross total resection (100% of the enhancing tumor volume) or near-total resection (>95% of the enhancing tumor volume) could be performed. The cases selected from the public datasets are referenced in Table C.3. For each patient, four MR structural weighted images — T1w, T2w, T2w-FLAIR, and post-T1w — were available. All the acquisitions of the private datasets were performed with IRB approval and informed written consent. See Table C.2 for details of the acquisition parameters.

Table C.1: Datasets used in this work. *BraTS2020* and *TCIA* are public datasets, whereas *Dataset22* and *Multicontrast Glioblastoma* are private datasets. The number of patients in each dataset is denoted by n. Age is shown as mean ± standard deviation. Survival is defined as the time in days from diagnosis to death (censored=0) or to the last date the patient was known to be alive (censored=1). The percentages of patients with survival less than 16 months (Survival < 16M) for the different datasets are also displayed (16 months or, equivalently, 480 days).

Dataset	**n**	**Age**	**Survival (IQR)**	**% Censored=1**	**Survival < 16M**
BraTS2020	119	62±12	374 (364)	0%	65.6 %
TCIA	34	60±10	521 (482)	5.9 %	58.8 %
Dataset22	22	65±10	451 (307)	22.7 %	59.1 %
Multicontrast Glioblastoma	24	57±13	552 (218)	29.2 %	54.2 %
Total	199	60±11	447 (346)	7.0 %	62.3 %

Table C.2: All MRI sessions are composed of four structural weighted images, namely, a T1w, a T2w, a T2w-FLAIR, and a post-T1w. Details about the scanner and the acquisition parameters are provided if available. The acquisition parameters are TE, TR, and TI.

Dataset	**Scanner**	**T1w**	**T2w**	**T2w-FLAIR**	**post-T1w**
BraTS2020	19 institutions	NA	NA	NA	NA
TCIA	8 institutions	TE = 2.75 - 19 ms TR = 352 - 3379 ms	TE = 15 - 120 ms TR = 700 - 6370 ms	TE = 34.6 - 155 ms TR = 6000 - 11000 ms	TE = 2.1 - 20 ms TR = 4.9 - 3285 ms
Dataset22	1.5T GE and 1.5T Philips	TE = 6.33 - 12 ms TR = 360 - 800 ms	TE = 99 - 110ms TR = 2680 - 8480 ms	TE = 120 - 127 ms TR = 6000 - 8000 ms TI = 2000 ms	TE = 2.56 ms TR = 7.96 ms
Multicontrast Glioblastoma	1.5T GE	TE = 1.83 ms TR = 5.98 ms	TE = 122 ms TR = 4162 ms	TE = 142 ms TR = 9350 ms TI = 2200 ms	TE = 2.18 ms TR = 6.85 ms

C.2 Radiomics system

C.2.1 Feature extraction

A total of five regions of interest (ROIs) were derived from the tumor segmentation. Three of them are directly defined in the segmentation, — enhancing tumor (ET), non-enhancing tumor (NET), edema (ED) —, and two more are constructed from them: tumor-core (TC), which is the union of ET and NET, and whole-tumor (WT), the union of the three initial regions. Moreover, 10 different filters are applied to the images. Wavelet filtering is applied using one decomposition level of the *db2 Daubechies orthogonal wavelet*, yielding eight decompositions; these are all possible combinations of applying either a high or a low pass filter in each of the three dimensions. Also, two *Laplacian of Gaussian* filters, are applied to emphasize areas of gray level change; this type of filter is controlled by the parameter sigma, which defines how coarse the emphasized texture should be. Two sigma values have been used (2 and 5). These features have shown improvements in the final prediction result (Chaddad et al., 2018).

From the previously mentioned five ROIs, four weightings and 11 images (both the original and the ten filtered images), a total of 117.088 features were extracted. MatLab was used to define and extract these features. Particularly, for texture features, the package presented by Dancheva et al. (2016) was employed. These features are defined as follows:

- Volume features (6 features): volume of the WT, the TC, each region (ET, NET, and ED) and the volume of the brain.
- Volume ratios (7 features): ratio of WT and the brain, ratio of ET and WT, ratio of NET and WT, ratio of ED and WT, ratio of ET and NET, ratio of ET and ED, and ratio of NET and ED.
- Morphological features (34 features): the percentage of the TC inside the cerebellum, brain stem, basal ganglia, and the parietal, occipital, frontal and temporal lobes. These measurements are obtained by segmenting the SRI-24 atlas with Freesurfer (Reuter et al., 2012), fusing the Freesurfer regions to create the preceding anatomical areas and taking advantage of the rigid registration of the different cases to the SRI-24 atlas to measure the presence of the tumor in those regions (see Figure C.1). These measurements are also calculated for the left and right hemispheres. Furthermore, the centroid of the tumor is determined and measurements of compactness, sphericity, the ratio volume to surface of the tumor, the TC surface area, and another sphericity measurement defined by Pérez-Beteta et al. (2018) are computed.
- First order, histogram-based, and texture features (68x4x5x11 features): 68 features are calculated over the whole ROIs. Calculated features are named in Table C.4.
- Extracted features from feature maps (58 x 8 x 4 x 5 x 11 features): Each feature from Table C.4 is computed on a 3 x 3 x 3 pixels block within each ROI to create different feature maps (see Figure C.2). In the Table C.4, the 20 histogram-based features are replaced by ten features derived from the frequency and probability values of the five-bin frequency and probability histograms constructed in the 3x3x3 blocks. As a consequence, the total number of features from which feature maps are derived is reduced from 68 to 58. Following that, the eight first-order features are calculated on these maps, yielding feature distribution measures along the feature maps. This method, which involves creating feature maps, may be able to improve the tumor heterogeneity characterization, which is a critical aspect of glioblastoma (Tirosh and Suvà, 2020).

Patient age is added to the group of the first three categories defined above, so this group consists of $6 + 7 + 34 + 1 = 48$ features. The total number of features can be calculated as the following summation:

$$(58 \times 8 + 68) \times 4 \times 5 \times 11 + 48 = 117.088,$$

being 4 the number of weighted images, 5 the number of ROIs and 11 the sum of the 10 filtered images —8 wavelet filters and 2 Laplacian of Gaussian filters— plus the original image.

C.2.2 Feature selection and model training

The radiomic system was trained following a 10-fold-within-5-fold nested cross-validation procedure. The method employed is the following:

- The total data was first divided in a training (175 patients) and testing set (24 patients). Outliers, defined as feature values above the 99^{th} percentile or below the 1^{th} percentile, were clipped and features were normalized by removing the mean and scaling to unit variance using the training set as reference.
- Five splits are performed over the training data (140 patients outer training and 35 patients outer validation). On each of these splits feature selection is executed.
- The outer training split is divided 10 times (125 patients for inner training, 15 patients for inner validation), and on these splits the training of 5 different machine learning models is carried out with the features selected in the outer split. Brier loss is calculated for each classifier and the best performer is chosen for each outer split.
- Each of the five models selected in the previous step is trained on its outer training data (140 patients) and validated on its outer validation data (35 patients). The model with the highest AUC on its validation data is the selected model.
- The selected model is finally trained on the training data (175 patients) and tested on the testing dataset (24 patients).

Feature selection in each outer loop is done following a three-step process:

1. First, the Spearman correlation between each feature and survival in days is calculated, and features that show a statistically significant correlation ($p < 0.05$) are retained. In addition, the Spearman correlation matrix between the statistically significant features is calculated, and the correlation values are used to eliminate those that are correlated with each other (correlation factor > 0.66), keeping the feature with the lowest p-value.
2. Second, the TuRF method was applied (Moore and White, 2007). TuRF adresses feature interaction iteratively utilizing the Relief method deriving feature statistics based on nearest neighbours. Features with the lowest scores are recursively eliminated. The number of features is narrowed to the top

100 with the best TuRF weights. In addition, age and 19 morphological and position information features, detailed in Table C.5, are reincorporated at this point. This type of features have been demonstrated as highly reproducible and able to improve model performance (Suter et al., 2020).

3. Finally, an information measure, maximum relevance-minimum redundancy (mRMR) (Ding and Peng, 2005), was utilized to obtain the final subset. This metric generates a ranking of feature groups by iterating in the cardinality of the groups. Specifically, mRMR is set to create fifty sets with cardinality one, 50 with cardinality two and so forth until cardinality 30. Each feature set that mRMR generates is tested on different classifiers in each outer split and, for the sets of the same cardinality, the one with the lowest Brier loss (in the inner splits) is selected. Calculating AUC with the validation data in the outer split, we select the best pair of classifier-cardinality of the feature set. The effective cardinality of feature sets, however, has been set smaller by applying the “one in ten rule” (Harrell Jr et al., 1984) so we have feature sets with cardinality less or equal to 17. Figure C.3, shows that only slight improvements in AUC are obtained when increasing the number of features beyond this threshold.

Four different models available in the Scikit-learn library (Pedregosa et al., 2011), namely, Naive Gaussian Bayes, LR, random forest, and SVM, as well as the XGB (Chen and Guestrin, 2016) model, were employed. The model and features selected for the experiment configurations I and II are defined in Table C.6. Model selected in this case was XGB. Feature importance, shown in Figure C.4, is calculated as the average gain of the splits on which the feature appears for all trees of the model (Chen and Guestrin, 2016). Furthermore, the features selected for the experiments with configuration III, both without T2w-FLAIR and without T2w, are listed in Tables C.7 and C.8. For these cases the models used were LR and SVM, respectively. Models’ hyperparameters during training are declared in Table C.9.

C.3 Supplementary material

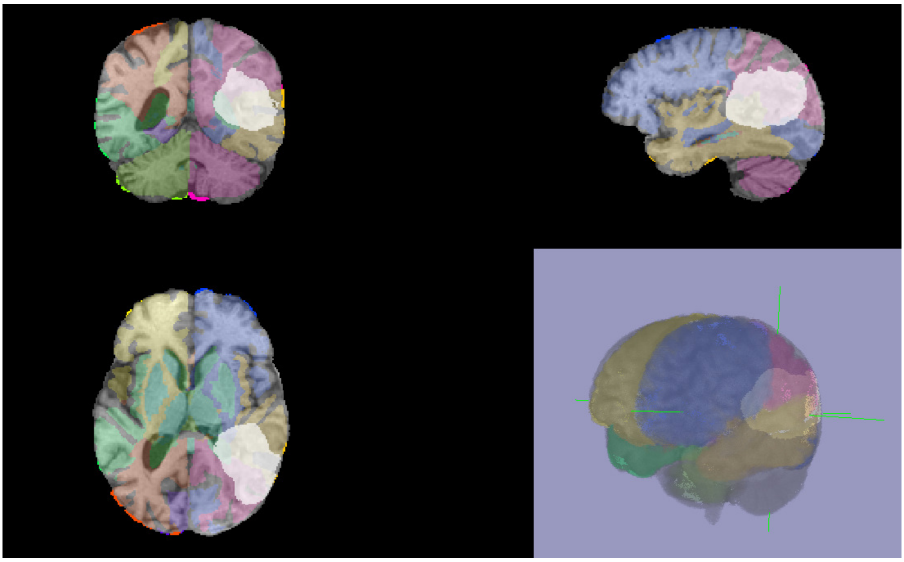


Figure C.1: Tumor-core (TC) presence in the different lobes of the brain for a representative training patient of *BraTS2020*. As seen in the image, the tumor is mainly located in the posterior part of the parietal and temporal left lobe. Features extracted are the percentage of the tumor in each region.

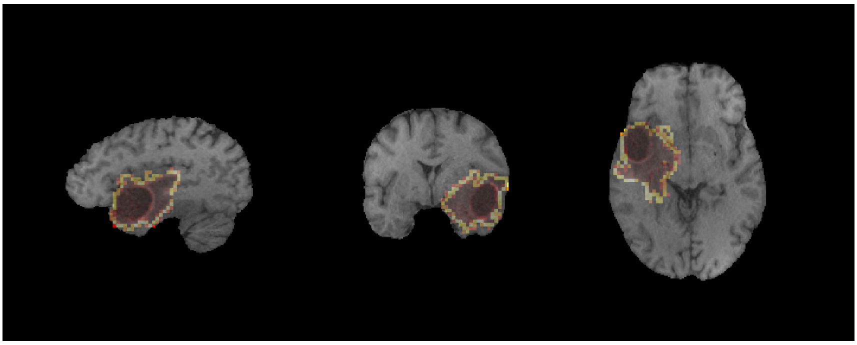


Figure C.2: Feature map standard deviation calculated over a block of 3x3x3 voxels in a representative training patient of the *BraTS2020* dataset.

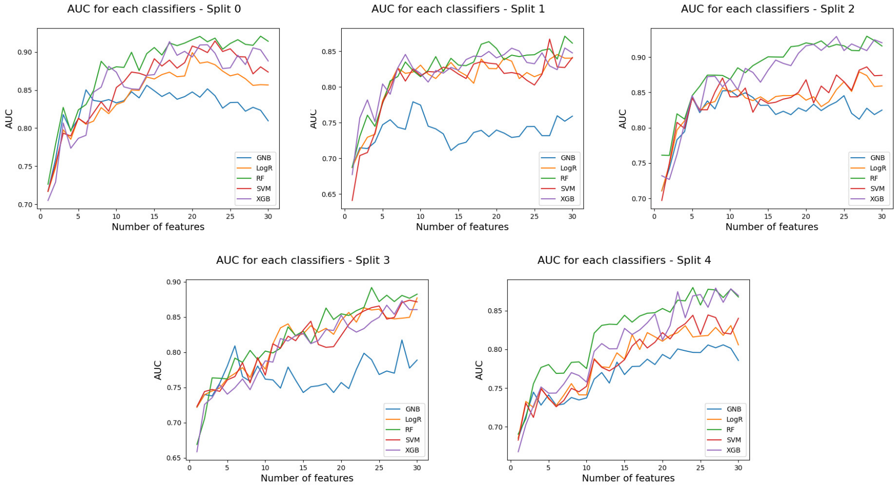


Figure C.3: AUC change when modifying the number of features for each outer split in the nested cross-validation scheme. The models evaluated are the best performers by AUC.

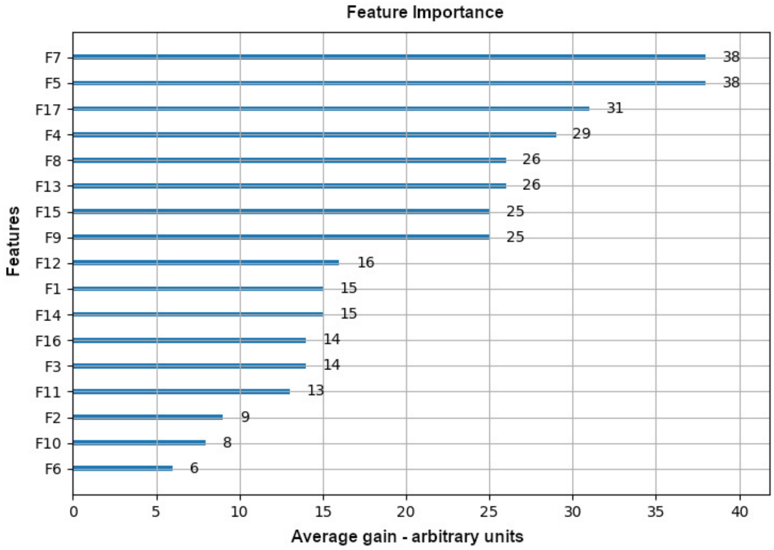


Figure C.4: Feature importance for the model selected. Importance is computed as the mean and standard deviation of accumulation of the impurity decrease (MDI) within each tree. Feature numeration follows Table C.6. T2w-FLAIR features are 9 and 11. T2w features are 10 and 14.

Table C.3: Cases selected from the public datasets *BraTS2020* and *Ivy-GAP*, *CPTAC*, and *TCGA* available through *TCIA*.

Dataset	Cases
BraTS2020	001, 002, 003, 004, 005, 006, 007, 009, 010, 012, 013, 014, 015, 016, 017, 018, 019, 021, 022, 023, 024, 028, 033, 034, 036, 037, 039, 042, 048, 049, 053, 054, 056, 057, 058, 059, 060, 061, 062, 064, 065, 066, 067, 068, 069, 070, 071, 072, 074, 076, 077, 078, 079, 080, 082, 083, 084, 086, 089, 090, 091, 093, 095, 096, 097, 098, 099, 100, 101, 102, 103, 104, 106, 107, 108, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 145, 147, 151, 152, 153, 154, 157, 337, 339, 341, 342, 345, 346, 347, 351, 356, 357, 358, 359, 360, 363, 366, 368, 369
Ivy-GAP	W1, W2, W5, W6, W8, W10, W11, W13, W19, W20, W22, W29, W32, W34, W35, W40, W43, W48, W54,
CPTAC	C3L_00278, C3L_00424, C3L_00528, C3L_00591, C3L_00677, C3L_02041, C3L_02504, C3L_03266, C3L_01505, C3L_00349, C3L_01327, C3L_02465, C3L_03727
TCGA	TCGA-14-1794, TCGA-14-1829

Table C.4: First order, histogram-based, and texture features. Texture features are extracted using 4 different texture metrics, neighbouring gray tone difference matrix (NGTDM), gray level run length matrix (GLRLM), gray level co-occurrence matrix (GLCM), and gray level size zone matrix (GLSZM).

Feature type	**Feature**	**Nº**
First order	Kurtosis, skewness, variance, standard deviation, median, mean minimum, and maximum	8
Histogram-based	Each bin of a 20 bins histogram	20
NGTDM	Coarseness, Contrast, Bussyness, Complexity, Strength	5
GLRLM	GLN, GLV, HGRE, LGRE, LRE, LRHGE, LRLGE, RLN, RLV, RP, SRE, SRHGE, SRLGE	13
GLCM	Energy, Contrast, Entropy, Homogeneity, Correlation, Variance, SumAverage, Auto Correlation, Dissimilarity	9
GLSZM	SZE, LZE, GLN, ZSN, ZP, LGZE, HGZE, SZLGE, SZHGE, LZLGE, LZHGE, GLV, ZLB	13

Table C.5: Features reintroduced in feature selection step 2.

Nº	**Feature**	**Feature definition**
1	Age	Age
2	rNET_ED	Volume ratio non-enhancing and tumor edema
3	rET_NET	Volume ratio enhancing tumor and non-enhancing tumor
4	rET_ED	Volume ratio enhancing tumor and edema
5	rNET_ED	Volume ratio non-enhancing tumor and edema
6	rED_WT	Volume ratio edema and whole tumor
7	rNET_WT	Volume ratio non-enhancing tumor and whole tumor
8	rET_WT	Volume ratio enhancing tumor and whole tumor
9	rNET_WT	Volume ratio non-enhancing tumor and whole tumor
10	rWT_B	Volume ratio whole tumor and brain
11	Morph_VolTumBasGang	Percentage of the tumor core volume in the basal ganglia measured by the rigid registration of the patients image data on the Freesurfer Segmentation of the SRI-24 atlas
12	Morph_mayAxisLegth	Mayor axis length of the tumor core
13	Morph_leastAxisLegth	Least axis length of the tumor core
14	Morph_coreArea	Area of the tumor core segmentation
15	Morph_sphericity	Sphericity calculated as $\frac{(36\times\pi\times coreVol^2)^{\frac{1}{3}}}{Morph\_coreArea}$
16	Morph_comp1	Compactness calculated as $\frac{coreVol}{\sqrt{(\pi)}\times\sqrt{(Morph\_coreArea^3)}}$
17	Morph_PB_Sr	Sphericity as specified in Pérez-Beteta et al. (2018) $6\times\sqrt{\pi}\times\frac{coreVol}{\sqrt{(Morph_coreArea^3)}}$
18	WT_cores	Number of WT focal points in the brain
19	TC_cores	Number of TC focal points in the brain
20	NET_cores	Number of non-enhancigng tumor focal points in the brain

Table C.6: Features selected for the model with four weighted images as input. The model selected in that case was extreme gradient boosting.

Nº	**Weighted image**	**ROI**	**Filter**	**Feature Map**	**Feature**
1	—	—	—	—	TC volume in basal ganglia
2	T1w	WT	Wavelet HLH	min	Variance
3	post-T1w	ET	Wavelet HHH	—	2nd bin 20 bins histogram
4	T1w	NET	LoG F2	GLSZM LZLGE	Variance
5	post-T1w	NET	Wavelet HHL	GLCM Homogeneity	Kurtosis
6	T1w	NET	Wavelet LLH	—	5th bin 20 bins histogram
7	post-T1w	TC	LoG F2	GLSZM LZHGE	Skewness
8	T1w	ED	Wavelet LHL	min	Stanadard Deviation
9	T2w-FLAIR	ET	Wavelet HLL	GLSZM SZHGE	Min
10	T2w	ET	Wavelet HLH	GLSZM ZSV	Mean
11	T2w-FLAIR	TC	Wavelet HLH	GLCM Dissimilarity	Skewness
12	T1w	ED	—	—	4th bin 20 bins histogram
13	T1w	WT	Wavelet LLH	GLRLM LGRE	Mean
14	T2w	ED	—	GLCM Contrast	Max
15	T1w	ET	LoG F2	GLRLM RLN	Skewness
16	T1w	ED	LoG F2	—	Skewness
17	T1w	ET	Wavelet HLH	NGTDM Busyness	Median

Table C.7: Features selected for the model with only three weighted images as inputs (without T2w-FLAIR). The model selected in that case was logistic regression.

Nº	**Weighted image**	**ROI**	**Filter**	**Feature Map**	**Feature**
1	post-T1w	ED	Wavelet LLL	Median	Mean
2	T1w	ET	Wavelet HLH	GLSZM SZE	Kurtosis
3	T1w	NET	Wavelet LLH	Min	Skewness
4	post-T1w	ED	Wavelet HLL	—	17th bin 20 bins histogram
5	T2w	ED	Wavelet HLH	GLSZM LZHGE	Variance
6	—	—	—	—	Volume ratio between ET and WT
7	T1w	ED	Wavelet HLL	GLSZM SZLGE	Mean
8	post-T1w	ED	Wavelet LLL	NGTDM Complexity	Standard Deviation
9	post-T1w	TC	—	GLSZM SZLGE	Skewness
10	T1w	WT	Wavelet LLH	GLRLM LRLGE	Skewness
11	post-T1w	ED	Wavelet HLL	GLRLM SRLGE	Min
12	T1w	TC	Wavelet HHL	Skewness	Median
13	T2w	TC	—	GLCM Dissimilarity	Max
14	T1w	ET	Wavelet LLL	Variance	Mean
15	T1w	TC	Wavelet LLH	GLSZM	Min
16	post-T1w	NET	Wavelet HHL	GLCM Homogeneity	Kurtosis

Table C.8: Features selected for the model with only three weighted images as input (without T2w). The model selected in that case was support vector machine.

Nº	**Weighted image**	**ROI**	**Filter**	**Feature Map**	**Feature**
1	T2w-FLAIR	TC	Wavelet HHL	GLSZM GLV	min
2	T2w-FLAIR	WT	Wavelet HLL	—	GLSZM SZHGE
3	T2w-FLAIR	ED	—	GLCM Contrast	Skewness
4	—	—	—	—	Age
5	T1w	ED	LoG F2	—	15$^{\text{th}}$ bin 20 bins histogram
6	T2w-FLAIR	NET	Wavelet HLL	GLRLM LGRE	Skewness
7	T1w	TC	Wavelet LLH	—	GLSZM SZLGE
8	T2w-FLAIR	WT	Wavelet LHL	Min	Kurtosis
9	T1w	TC	—	GLRLM LRHGE	Standard deviation
10	T1w	ED	Wavelet LLL	—	3$^{\text{rd}}$ bin 20 bins histogram
11	—	—	—	—	NET focal points
12	T1w	ED	Wavelet LHL	Mean	Median
13	T2w-FLAIR	WT	Wavelet HHH	—	GLSZM LGZE
14	T2w-FLAIR	ET	Wavelet HHL	GLCM SumAverage	Min
15	T1w	WT	Wavelet LHH	GLRLM LRHGE	Median
16	T1w	ET	LoG F2	—	GLRLM HGRE

Table C.9: Models' hyperparameters.

Model	**Hyperparameters**
Logistic Regression	penalty='l2', *, dual=False, tol=0.0001, C=1.0, fit_intercept=True, intercept_scaling=1, solver='lbfgs', max_iter=100, multi_class='auto', verbose=0, warm_start=False,
Naive Gaussian Bayes	*, priors=None, var_smoothing=1e-09
Random Forrest	n_estimators=500, *, criterion='entropy', max_depth=4 min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features='auto', min_impurity_decrease=0.0, bootstrap=True, oob_score=False, n_jobs=None, verbose=0, warm_start=False, ccp_alpha=0.0,
Support Vector Machine	*, C=1.0, kernel='linear', degree=3, gamma='scale', coef0=0.0, shrinking=True, probability=False, tol=0.0001, cache_size=200, verbose=False, max_iter=- 1, decision_function_shape='ovr', break_ties=False
Extreme Gradient Boosting	objective: 'binary:logistic', use_label_encoder: False, importance_type: 'gain', n_estimators: 100, verbosity: 0,

D

Voxelwise statistical classifications for tumor T1w-enhancement prediction

Nunez-Gonzalez et al. (2022) proposed a method for voxel-wise classification of normal/abnormal tissue and tissue with and without T1w-enhancement from only pre-contrast MAGiC parametric maps. In this study, the authors performed the voxel-wise classifications using a receiver operating characteristic (ROC) curve analysis. The authors considered three classification problems:

C-I) ABN versus nWM.

C-II) T1e versus the union of nWM and T2h.

C-III) T1e versus T2h (only inside ABN).

For each classification problem, they defined four voxel-wise metrics:

a) **T1** values measured in miliseconds.

b) **T2** values measured in miliseconds.

c) ***normT1T2*** (i.e., the Euclidean norm of the T1 and T2 values).

d) ***normlog*** (i.e., the Euclidean norm of the logarithm of T1 and T2 values).

For each classification problem the metric with the highest AUC was selected. The thresholds for optimal classifications were calculated as the highest Youden's index of the ROC curve. In a second step, the classification thresholds were applied to the aforementioned metrics inside the white-matter-mask in order to obtain the different regions. The selected regions were compared with the ground-truth segmentations obtained with the deep learning tool HD-GLIO fed by the T1w, T2w, T2w-FLAIR, and post-T1w (Isensee et al., 2019; Kickingereder et al., 2019).

Thus, the resulting selected metrics and their thresholds were: *normlog* with threshold 8.44 for C-I, *normT1T2* with threshold 1344 ms for C-II, and *normT1T2* with threshold 1512 ms for C-III. The authors showed that The ROC analysis including PD did not improve the AUC in any of the cases. Consequently, they excluded the PD values in the rest of the analysis.

Bibliography

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., et al., 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467 .

Adrion, W.R., 1993. Research methodology in software engineering, in: Summary of the Dagstuhl Workshop on Future Directions in Software Engineering. Ed. Tichy, Habermann, and Prechelt, ACM Software Engineering Notes, SIGSoft, pp. 36–37.

Armanious, K., Jiang, C., Fischer, M., Küstner, T., Hepp, T., Nikolaou, K., Gatidis, S., Yang, B., 2020. MedGAN: Medical image translation using GANs. Computerized medical imaging and graphics 79, 101684. doi:`10.1016/j.compmedimag.2019.101684`.

Arnold, T.C., Freeman, C.W., Litt, B., Stein, J.M., 2023. Low-field MRI: clinical promise and challenges. Journal of Magnetic Resonance Imaging 57, 25–44. doi:`10.1002/jmri.28408`.

Arvanitis, C.D., Ferraro, G.B., Jain, R.K., 2020. The blood–brain barrier and blood–tumour barrier in brain tumours and metastases. Nature Reviews Cancer 20, 26–41. doi:`10.1038/s41568-019-0205-x`.

Badve, C., Yu, A., Dastmalchian, S., Rogers, M., Ma, D., Jiang, Y., Margevicius, S., Pahwa, S., Lu, Z., Schluchter, M., et al., 2017. Mr fingerprinting of adult brain tumors: initial experience. American Journal of Neuroradiology 38, 492–499. doi:`10.3174/ajnr.A5035`.

Bakas, S., Sako, C., Akbari, H., Bilello, M., Sotiras, A., Shukla, G., Rudie, J.D., Santamaría, N.F., Kazerooni, A.F., Pati, S., et al., 2022. The University of Pennsylvania glioblastoma (UPenn-GBM) cohort: Advanced MRI, clinical, genomics, & radiomics. Scientific data 9, 453. doi:`10.1038/s41597-022-01560-7`.

Barbieri, S., Gurney-Champion, O.J., Klaassen, R., Thoeny, H.C., 2020. Deep learning how to fit an intravoxel incoherent motion model to diffusion-weighted MRI. Magnetic Resonance in Medicine 83, 312–321. doi:`10.1002/mrm.27910`.

Bernstein, M.A., King, K.F., Zhou, X.J., 2004. Handbook of MRI pulse sequences. Elsevier.

Bittoun, J., Taquin, J., Sauzade, M., 1984. A computer algorithm for the simulation of any nuclear magnetic resonance (NMR) imaging method. Magnetic Resonance Imaging 2, 113–120. doi:`10.1016/0730-725X(84)90065-1`.

Bloch, F., 1946. Nuclear induction. Physical review 70, 460. doi:`10.1103/PhysRev.70.460`.

Blystad, I., Warntjes, J., Smedby, Ö., Lundberg, P., Larsson, E.M., Tisell, A., 2020. Quantitative MRI using relaxometry in malignant gliomas detects contrast enhancement in peritumoral oedema. Scientific reports 10, 1–9. doi:`10.1038/s41598-020-75105-6`.

Blystad, I., Warntjes, J.B.M., Smedby, O., Landtblom, A.M., Lundberg, P., Larsson, E.M., 2012. Synthetic MRI of the brain in a clinical setting. Acta Radiologica 53, 1158–1163. doi:`10.1258/ar.2012.120195`.

Blystad, I., Warntjes, J.M., Smedby, Ö., Lundberg, P., Larsson, E.M., Tisell, A., 2017. Quantitative MRI for analysis of peritumoral edema in malignant gliomas. PLoS One 12, e0177135. doi:`10.1371/journal.pone.0177135`.

Bobman, S.A., Riederer, S.J., Lee, J.N., Suddarth, S.A., Wang, H.Z., Drayer, B.P., MacFall, J.R., 1985. Cerebral magnetic resonance image synthesis. American journal of neuroradiology 6, 265–269.

Bojorquez, J.Z., Bricq, S., Acquitter, C., Brunotte, F., Walker, P.M., Lalande, A., 2017. What are normal relaxation times of tissues at 3 T? Magnetic Resonance Imaging 35, 69–80. doi:`10.1016/j.mri.2016.08.021`.

Brekenfeld, C., Foert, E., Hundt, W., Kenn, W., Lodeann, K.P., GEHL, H.B., 2001. Enhancement of cerebral diseases: How much contrast agent is enough?: Comparison of 0.1, 0.2, and 0.3 mmol/kg gadoteridol at 0.2 T with 0.1 mmol/kg gadoteridol at 1.5 T. Investigative Radiology 36, 266–275. doi:`10.1097/00004424-200105000-00004`.

Brown, T.J., Brennan, M.C., Li, M., Church, E.W., Brandmeir, N.J., Rakszawski, K.L., Patel, A.S., Rizk, E.B., Suki, D., Sawaya, R., et al., 2016. Association of the extent of resection with survival in glioblastoma: a systematic review and meta-analysis. JAMA oncology 2, 1460–1469. doi:`10.1001/jamaoncol.2016.1373`.

Cao, Z., Oh, S., Sica, C.T., McGarrity, J.M., Horan, T., Luo, W., Collins, C.M., 2014. Bloch-based MRI system simulator considering realistic electromagnetic fields for calculation of signal, noise, and specific absorption rate. Magnetic resonance in medicine 72, 237–247. doi:`10.1002/mrm.24907`.

Castillo-Passi, C., Coronado, R., Varela-Mattatall, G., Alberola-López, C., Botnar, R., Irarrazaval, P., 2023. KomaMRI.jl: An open-source framework for general MRI simulations with GPU acceleration. Magnetic Resonance in Medicine 90, 329–342. doi:`10.1002/mrm.29635`.

Chaddad, A., Daniel, P., Niazi, T., 2018. Radiomics evaluation of histological heterogeneity using multiscale textures derived from 3D wavelet transformation of multispectral images. Front Oncol 8, 96. doi:`10.3389/fonc.2018.00096`.

Chartsias, A., Joyce, T., Giuffrida, M.V., Tsaftaris, S.A., 2017. Multimodal MR synthesis via modality-invariant latent representation. IEEE Transactions on Medical Imaging 37, 803–814. doi:`10.1109/TMI.2017.2764326`.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, San Francisco, CA, USA. pp. 785–794.

Cheng, C.C., Preiswerk, F., Madore, B., 2020. Multi-pathway multi-echo acquisition and neural contrast translation to generate a variety of quantitative and qualitative image contrasts. Magnetic Resonance in Medicine 83, 2310–2321. doi:`10.1002/mrm.28077`.

Chiou, E., Valindria, V., Giganti, F., Punwani, S., Kokkinos, I., Panagiotaki, E., 2021. Synthesizing VERDICT maps from standard DWI data using GANs, in: Computational Diffusion MRI: 12th International Workshop, CDMRI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, October 1, 2021, Proceedings 12, Springer. pp. 58–67.

Clare, S., Jezzard, P., 2001. Rapid T1 mapping using multislice echo planar imaging. Magnetic Resonance in Medicine 45, 630–634. doi:`10.1002/mrm.1085`.

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. J Digit Imaging 26, 1045–1057. doi:`10.1007/s10278-013-9622-7`.

Cocosco, C.A., Kollokian, V., Kwan, R.K.S., Pike, G.B., Evans, A.C., 1997. BrainWeb: Online Interface to a 3D MRI Simulated Brain Database. NeuroImage 5, 425.

Coelho, S., Liao, Y., Szczepankiewicz, F., Veraart, J., Chung, S., Lui, Y.W., Novikov, D.S., Fieremans, E., 2024. Assessment of precision and accuracy of brain white matter microstructure using combined diffusion MRI and relaxometry. arXiv preprint arXiv:2402.17175 .

Conlon, P., Trimble, M., Rogers, D., Callicott, C., 1988. Magnetic resonance imaging in epilepsy: a controlled study. Epilepsy research 2, 37–43. doi:`10.1016/0920-1211(88)90008-3`.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. Mathematics of control, signals and systems 2, 303–314. doi:`10.1007/BF02551274`.

Dai, X., Lei, Y., Fu, Y., Curran, W.J., Liu, T., Mao, H., Yang, X., 2020. Multimodal MRI synthesis using unified generative adversarial networks. Med Phys 47, 6343–6354. doi:`10.1002/mp.14539`.

Dalmaz, O., Yurt, M., Çukur, T., 2022. ResViT: residual vision transformers for multimodal medical image synthesis. IEEE Transactions on Medical Imaging 41, 2598–2614. doi:`10.1109/TMI.2022.3167808`.

Dancheva, Z., Bochev, P., Chaushev, B., Yordanova, T., Klisarova, A., 2016. Dual-time point 18FDG-PET/CT imaging may be useful in assessing local recurrent disease in high grade bone and soft tissue sarcoma. Nucl Med Rev 19, 22–27. doi:`10.5603/NMR.2016.0005`.

Dar, S.U., Yurt, M., Karacan, L., Erdem, A., Erdem, E., Çukur, T., 2019. Image synthesis in multi-contrast MRI with conditional generative adversarial networks. IEEE Transactions on Medical Imaging 38, 2375–2388. doi:`10.1109/TMI.2019.2901750`.

Deoni, S.C., 2007. High-resolution T1 mapping of the brain at 3T with driven equilibrium single pulse observation of T1 with high-speed incorporation of RF field inhomogeneities (DESPOT1-HIFI). Journal of Magnetic Resonance Imaging 26, 1106–1111. doi:`10.1002/jmri.21130`.

Deoni, S.C., 2010. Quantitative relaxometry of the brain. Topics in magnetic resonance imaging: TMRI 21, 101. doi:`10.1097/RMR.0b013e31821e56d8`.

Deoni, S.C., Peters, T.M., Rutt, B.K., 2005. High-resolution T1 and T2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2. Magnetic Resonance in Medicine 53, 237–241. doi:`10.1002/mrm.20314`.

Deoni, S.C., Ward, H.A., Peters, T.M., Rutt, B.K., 2004. Rapid T2 estimation with phase-cycled variable nutation steady-state free precession. Magnetic Resonance in Medicine 52, 435–439. doi:`10.1002/mrm.20159`.

Desai, N.K., Runge, V.M., 2003. Contrast use at low field: a review. Topics in Magnetic Resonance Imaging 14, 360–364. doi:10.1097/00002142-200310000-00002.

DeWitt, L.D., Kistler, J.P., Miller, D.C., Richardson Jr, E.P., Buonanno, F.S., 1987. NMR-neuropathologic correlation in stroke. Stroke 18, 342–351. doi:10.1161/01.STR.18.2.342.

Ding, C., Peng, H., 2005. Minimum redundancy feature selection from microarray gene expression data. J Bioinform Comput Biol 3, 185–205. doi:10.1142/S0219720005001004.

Ehrlich, H., Rivera, M., 2021. AxonNet: A self-supervised Deep Neural Network for Intravoxel Structure Estimation from DW-MRI. arXiv preprint arXiv:2103.11006 .

Ellingson, B.M., Bendszus, M., Boxerman, J., Barboriak, D., Erickson, B.J., Smits, M., Nelson, S.J., Gerstner, E., Alexander, B., Goldmacher, G., et al., 2015a. Consensus recommendations for a standardized Brain Tumor Imaging Protocol in clinical trials. Neuro-oncology 17, 1188–1198. doi:10.1093/neuonc/nov095.

Ellingson, B.M., Lai, A., Nguyen, H.N., Nghiemphu, P.L., Pope, W.B., Cloughesy, T.F., 2015b. Quantification of nonenhancing tumor burden in gliomas using effective T2 maps derived from dual-echo turbo spin-echo MRI. Clinical Cancer Research 21, 4373–4383. doi:10.1158/1078-0432.CCR-14-2862.

Errante, Y., Cirimele, V., Mallio, C.A., Di Lazzaro, V., Zobel, B.B., Quattrocchi, C.C., 2014. Progressive increase of T1 signal intensity of the dentate nucleus on unenhanced magnetic resonance images is associated with cumulative doses of intravenously administered gadodiamide in patients with normal renal function, suggesting dechelation. Investigative radiology 49, 685–690. doi:10.1097/RLI.0000000000000072.

Evans, A.C., Janke, A.L., Collins, D.L., Baillet, S., 2012. Brain templates and atlases. Neuroimage 62, 911–922. doi:10.1016/j.neuroimage.2012.01.024.

Fokkinga, E., Hernandez-Tamames, J.A., Ianus, A., Nilsson, M., Tax, C.M., Perez-Lopez, R., Grussu, F., 2023. Advanced diffusion-weighted MRI for cancer microstructure assessment in body imaging, and its relationship with histology. Journal of Magnetic Resonance Imaging doi:10.1002/jmri.29144.

Forghani, R., 2016. Adverse effects of gadolinium-based contrast agents: changes in practice patterns. Topics in Magnetic Resonance Imaging 25, 163–169. doi:10.1097/RMR.0000000000000095.

Fraum, T.J., Ludwig, D.R., Bashir, M.R., Fowler, K.J., 2017. Gadolinium-based contrast agents: a comprehensive risk assessment. Journal of Magnetic Resonance Imaging 46, 338–353. doi:10.1002/jmri.25625.

Gelman, N., Gorell, J.M., Barker, P.B., Savage, R.M., Spickler, E.M., Windham, J.P., Knight, R.A., 1999. MR imaging of human brain at 3.0 T: preliminary report on transverse relaxation rates and relation to estimated iron content. Radiology 210, 759–767. doi:10.1148/radiology.210.3.r99fe41759.

Gilard, V., Tebani, A., Dabaj, I., Laquerrière, A., Fontanilles, M., Derrey, S., Marret, S., Bekri, S., 2021. Diagnosis and management of glioblastoma: A comprehensive perspective. Journal of Personalized Medicine 11, 258. doi:10.3390/jpm11040258.

Glass, R.L., 1995. A structure-based critique of contemporary computing research. Journal of Systems and Software 28, 3–7. doi:10.1016/0164-1212(94)00077-Z.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative Adversarial Nets, in: Advances in Neural Information Processing Systems.

Granata, V., Cascella, M., Fusco, R., Catalano, O., Filice, S., Schiavone, V., Izzo, F., Cuomo, A., Petrillo, A., et al., 2016. Immediate adverse reactions to gadolinium-based MR contrast media: a retrospective analysis on 10,608 examinations. BioMed research international 2016. doi:`10.1155/2016/3918292`.

Grobner, T., 2006. Gadolinium–a specific trigger for the development of nephrogenic fibrosing dermopathy and nephrogenic systemic fibrosis? Nephrology Dialysis Transplantation 21, 1104–1108. doi:`10.1093/ndt/gfl294`.

Gulani, V., Calamante, F., Shellock, F.G., Kanal, E., Reeder, S.B., 2017. Gadolinium deposition in the brain: summary of evidence and recommendations. The Lancet Neurology 16, 564–570. doi:`10.1016/S1474-4422(17)30158-8`.

Gulani, V., Schmitt, P., Griswold, M.A., Webb, A.G., Jakob, P.M., 2004. Towards a single-sequence neurologic magnetic resonance imaging examination: multiple-contrast images from an IR TrueFISP experiment. Investigative radiology 39, 767–774. doi:`10.1097/00004424-200412000-00008`.

Gulani, V., Seiberlich, N., 2020. Quantitative MRI: Rationale and Challenges. doi:`10.1016/B978-0-12-817057-1.00001-9`.

Hagiwara, A., Hori, M., Cohen-Adad, J., Nakazawa, M., Suzuki, Y., Kasahara, A., Horita, M., Haruyama, T., Andica, C., Maekawa, T., et al., 2019a. Linearity, bias, intrascanner repeatability, and interscanner reproducibility of quantitative multidynamic multiecho sequence for rapid simultaneous relaxometry at 3 T: a validation study with a standardized phantom and healthy controls. Investigative Radiology 54, 39–47. doi:`0.1097/RLI.0000000000000510`.

Hagiwara, A., Otsuka, Y., Hori, M., Tachibana, Y., Yokoyama, K., Fujita, S., Andica, C., Kamagata, K., Irie, R., Koshino, S., et al., 2019b. Improving the quality of synthetic FLAIR images with deep learning using a conditional generative adversarial network for pixel-by-pixel image translation. American Journal of Neuroradiology 40, 224–230. doi:`10.3174/ajnr.A5927`.

Hagiwara, A., Warntjes, M., Hori, M., Andica, C., Nakazawa, M., Kumamaru, K.K., Abe, O., Aoki, S., 2017. SyMRI of the brain: rapid quantification of relaxation rates and proton density, with synthetic MRI, automatic brain segmentation, and myelin measurement. Investigative radiology 52, 647–657. doi:`10.1097/RLI.0000000000000365`.

Harrell Jr, F.E., Lee, K.L., Califf, R.M., Pryor, D.B., Rosati, R.A., 1984. Regression modelling strategies for improved prognostic prediction. Stat Med 3, 143–152. doi:`10.1002/sim.4780030207`.

Hattingen, E., Jurcoane, A., Daneshvar, K., Pilatus, U., Mittelbronn, M., Steinbach, J.P., Bähr, O., 2013. Quantitative T2 mapping of recurrent glioblastoma under bevacizumab improves monitoring for non-enhancing tumor progression and predicts overall survival. Neuro-Oncology 15, 1395–1404. doi:`10.1093/neuonc/not105`.

Hattingen, E., Müller, A., Jurcoane, A., Mädler, B., Ditter, P., Schild, H., Herrlinger, U., Glas, M., Kebir, S., 2017. Value of quantitative magnetic resonance imaging T1-relaxometry in predicting contrast-enhancement in glioblastoma patients. Oncotarget 8, 53542. doi:`10.18632/oncotarget.18612`.

Holland, E.C., 2001. Progenitor cells and glioma formation. Current opinion in neurology 14, 683–688. doi:`10.1097/00019052-200112000-00002`.

Horbinski, C., Berger, T., Packer, R.J., Wen, P.Y., 2022. Clinical implications of the 2021 edition of the WHO classification of central nervous system tumours. Nature Reviews Neurology 18, 515–529. doi:`10.1038/s41582-022-00679-w`.

Hori, M., Hagiwara, A., Goto, M., Wada, A., Aoki, S., 2021. Low-field magnetic resonance imaging: its history and renaissance. Investigative radiology 56, 669–679. doi:`10.1097/RLI.0000000000000810`.

Hu, L., Zhou, D.w., Zha, Y.f., Li, L., He, H., Xu, W.h., Qian, L., Zhang, Y.k., Fu, C.x., Hu, H., et al., 2021. Synthesizing High-b-Value Diffusion–weighted Imaging of the Prostate using Generative Adversarial Networks. Radiology: Artificial Intelligence 3, e200237. doi:`10.1148/ryai.2021200237`.

Inoue, K., Fukushi, M., Furukawa, A., Sahoo, S.K., Veerasamy, N., Ichimura, K., Kasahara, S., Ichihara, M., Tsukada, M., Torii, M., et al., 2020. Impact on gadolinium anomaly in river waters in Tokyo related to the increased number of MRI devices in use. Marine Pollution Bulletin 154, 111148. doi:`10.1016/j.marpolbul.2020.111148`.

Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H., 2021. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods 18, 203–211. doi:`10.1038/s41592-020-01008-z`.

Isensee, F., Petersen, J., Kohl, S.A., Jäger, P.F., Maier-Hein, K.H., 2019. nnu-net: Breaking the spell on successful medical image segmentation. arXiv preprint arXiv:1904.08128 1, 2.

Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks, in: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1125–1134.

Iturri-Clavero, F., Galbarriatu-Gutierrez, L., Gonzalez-Uriarte, A., Tamayo-Medel, G., de Orte, K., Martinez-Ruiz, A., Castellon-Larios, K., Bergese, S., 2016. “low-field” intraoperative MRI: a new scenario, a new adaptation. Clinical Radiology 71, 1193–1198. doi:`10.1016/j.crad.2016.07.003`.

Jais, I.K.M., Ismail, A.R., Nisa, S.Q., 2019. Adam optimization algorithm for wide and deep neural network. Knowledge Engineering and Data Science 2, 41–46. doi:`10.17977/um018v2i12019p41-46`.

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. FSL. Neuroimage 62, 782–790. doi:`10.1016/j.neuroimage.2011.09.015`.

Ji, S., Yang, D., Lee, J., Choi, S.H., Kim, H., Kang, K.M., 2020. Synthetic MRI: Technologies and Applications in Neuroradiology. Journal of Magnetic Resonance Imaging doi:`10.1002/jmri.27440`.

Kanda, T., Ishii, K., Kawaguchi, H., Kitajima, K., Takenaka, D., 2014. High signal intensity in the dentate nucleus and globus pallidus on unenhanced T1-weighted MR images: relationship with increasing cumulative dose of a gadolinium-based contrast material. Radiology 270, 834–841. doi:`10.1148/radiol.13131669`.

Kickingereder, P., Isensee, F., Tursunova, I., Petersen, J., Neuberger, U., Bonekamp, D., Brugnara, G., Schell, M., Kessler, T., Foltyn, M., et al., 2019. Automated quantitative tumour response assessment of MRI in neuro-oncology with artificial neural networks:

a multicentre, retrospective study. The Lancet Oncology 20, 728–740. doi:`10.1016/S1470-2045(19)30098-1`.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 .

Kjær, L., Thomsen, C., Gjerris, F., Mosdal, B., Henriksen, O., 1991. Tissue characterization of intracranial tumors by MR imaging: in vivo evaluation of T1-and T2-relaxation behavior at 1.5T. Acta Radiologica 32, 498–504. doi:`10.3109/02841859109177614`.

Kleesiek, J., Morshuis, J.N., Isensee, F., Deike-Hofmann, K., Paech, D., Kickingereder, P., Köthe, U., Rother, C., Forsting, M., Wick, W., et al., 2019. Can virtual contrast enhancement in brain MRI replace gadolinium?: a feasibility study. Invest Radiol 54, 653–660. doi:`10.1097/RLI.0000000000000583`.

Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 15, 155–163. doi:`10.1016/j.jcm.2016.02.012`.

Kurki, T., Komu, M., 1995. Spin-lattice relaxation and magnetization transfer in intracranial tumors in vivo: effects of Gd-DTPA on relaxation parameters. Magnetic resonance imaging 13, 379–385. doi:`10.1016/0730-725X(94)00126-N`.

Lacroix, M., Abi-Said, D., Fourney, D.R., Gokaslan, Z.L., Shi, W., DeMonte, F., Lang, F.F., McCutcheon, I.E., Hassenbusch, S.J., Holland, E., et al., 2001. A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival. Journal of neurosurgery 95, 190–198. doi:`10.3171/jns.2001.95.2.0190`.

Larsson, H., Frederiksen, J., Petersen, J., Nordenbo, A., Zeeberg, I., Henriksen, O., Olesen, J., 1989. Assessment of demyelination, edema, and gliosis by in vivo determination of T1 and T2 in the brain of patients with acute attack of multiple sclerosis. Magnetic Resonance in Medicine 11, 337–348. doi:`10.1002/mrm.1910110308`.

Lauterbur, P.C., 1973. Image formation by induced local interactions: examples employing nuclear magnetic resonance. nature 242, 190–191. doi:`10.1038/242190a0`.

Lescher, S., Jurcoane, A., Veit, A., Bähr, O., Deichmann, R., Hattingen, E., 2015. Quantitative T1 and T2 mapping in recurrent glioblastomas under bevacizumab: earlier detection of tumor progression compared to conventional MRI. Neuroradiology 57, 11–20. doi:`10.1007/s00234-014-1445-9`.

Li, X., Wang, L., Liu, H., Ma, B., Chu, L., Dong, X., Zeng, D., Che, T., Jiang, X., Wang, W., et al., 2023. Syn_SegNet: A joint deep neural network for ultrahigh-field 7T MRI synthesis and hippocampal subfield segmentation in routine 3T MRI. IEEE Journal of Biomedical and Health Informatics doi:`10.1109/JBHI.2023.3305377`.

Liberman, G., Louzoun, Y., Ben Bashat, D., 2014. T1 mapping using variable flip angle SPGR data with flip angle correction. Journal of Magnetic Resonance Imaging 40, 171–180. doi:`10.1002/jmri.24373`.

Liu, J., Pasumarthi, S., Duffy, B., Gong, E., Datta, K., Zaharchuk, G., 2023. One model to synthesize them all: Multi-contrast multi-scale transformer for missing data imputation. IEEE Transactions on Medical Imaging doi:`10.1109/TMI.2023.3261707`.

Liu, Y., Chen, A., Shi, H., Huang, S., Zheng, W., Liu, Z., Zhang, Q., Yang, X., 2021. CT synthesis from MRI using multi-cycle GAN for head-and-neck radiation therapy.

Computerized medical imaging and graphics 91, 101953. doi:`10.1016/j.compmedimag.2021.101953`.

Lohrke, J., Frenzel, T., Endrikat, J., Alves, F.C., Grist, T.M., Law, M., Lee, J.M., Leiner, T., Li, K.C., Nikolaou, K., et al., 2016. 25 years of contrast-enhanced MRI: developments, current challenges and future perspectives. Advances in therapy 33, 1–28. doi:`10.1007/s12325-015-0275-4`.

Lu, H., Nagae-Poetscher, L.M., Golay, X., Lin, D., Pomper, M., Van Zijl, P.C., 2005. Routine clinical brain MRI sequences for use at 3.0 Tesla. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine 22, 13–22. doi:`10.1002/jmri.20356`.

Ma, D., Gulani, V., Seiberlich, N., Liu, K., Sunshine, J.L., Duerk, J.L., Griswold, M.A., 2013. Magnetic resonance fingerprinting. Nature 495, 187–192. doi:`10.1038/nature11971`.

Maas, A.L., Hannun, A.Y., Ng, A.Y., et al., 2013. Rectifier nonlinearities improve neural network acoustic models, in: Proc. icml, Atlanta, GA. p. 3.

Maggioni, M., Foi, A., 2012. Nonlocal transform-domain denoising of volumetric data with groupwise adaptive variance estimation, International Society for Optics and Photonics, Burlingame, CA, USA. p. 82960O.

McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics 5, 115–133. doi:`10.1007/BF02478259`.

McNamara, C., Mankad, K., Thust, S., Dixon, L., Limback-Stanic, C., D'Arco, F., Jacques, T.S., Löbel, U., 2022. 2021 WHO classification of tumours of the central nervous system: a review for the neuroradiologist. Neuroradiology 64, 1919–1950. doi:`10.1007/s00234-022-03008-6`.

McRobbie, D.W., Moore, E.A., Graves, M.J., Prince, M.R., 2017. MRI from Picture to Proton. Cambridge university press.

Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al., 2014. The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans Med Imaging 34, 1993–2024. doi:`10.1109/TMI.2014.2377694`.

Moore, J.H., White, B.C., 2007. Tuning ReliefF for genome-wide genetic analysis, Springer, Valencia, Spain. pp. 166–175. doi:`10.1007/978-3-540-71783-6_16`.

Moya-Sáez, E., Peña-Nogales, Ó., de Luis-García, R., Alberola-López, C., 2021. A deep learning approach for synthetic MRI based on two routine sequences and training with synthetic data. Computer Methods and Programs in Biomedicine 210, 106371. doi:`10.1016/j.cmpb.2021.106371`.

Müller, A., Jurcoane, A., Kebir, S., Ditter, P., Schrader, F., Herrlinger, U., Tzaridis, T., Mädler, B., Schild, H.H., Glas, M., et al., 2017. Quantitative T1-mapping detects cloudy-enhancing tumor compartments predicting outcome of patients with glioblastoma. Cancer medicine 6, 89–99. doi:`10.1002/cam4.966`.

Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines, in: Proceedings of the 27th international conference on machine learning (ICML-10), pp. 807–814.

Naruse, S., Horikawa, Y., Tanaka, C., Hirakawa, K., Nishikawa, H., Yoshizaki, K., 1986. Significance of proton relaxation time measurement in brain edema, cerebral infarction and brain tumors. Magnetic resonance imaging 4, 293–304. doi:`10.1016/0730-725X(86)91039-8`.

Nunez-Gonzalez, L., van Garderen, K.A., Smits, M., Jaspers, J., Romero, A.M., Poot, D.H., Hernandez-Tamames, J.A., 2022. Pre-contrast MAGiC in treated gliomas: a pilot study of quantitative MRI. Scientific Reports 12, 21820. doi:`10.1038/s41598-022-24276-5`.

Oh, J., Cha, S., Aiken, A.H., Han, E.T., Crane, J.C., Stainsby, J.A., Wright, G.A., Dillon, W.P., Nelson, S.J., 2005. Quantitative apparent diffusion coefficients and T2 relaxation times in characterizing contrast enhancing brain tumors and regions of peritumoral edema. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine 21, 701–708. doi:`10.1002/jmri.20335`.

Ostrom, Q.T., Patil, N., Cioffi, G., Waite, K., Kruchko, C., Barnholtz-Sloan, J.S., 2020. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2013–2017. Neuro-oncology 22, iv1–iv96. doi:`10.1093/neuonc/noaa200`.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al., 2011. Scikit-learn: Machine learning in Python. J Mach Learn Res 12, 2825–2830.

Peña-Nogales, Ó., Ellmore, T.M., de Luis-García, R., Suescun, J., Schiess, M.C., Giancardo, L., 2019. Longitudinal connectomes as a candidate progression marker for prodromal Parkinson's disease. Frontiers in Neuroscience 12, 967. doi:`10.3389/fnins.2018.00967`.

Pérez-Beteta, J., Molina-García, D., Ortiz-Alhambra, J.A., Fernández-Romero, A., Luque, B., Arregui, E., Calvo, M., Borrás, J.M., Meléndez, B., Rodriguez de Lope, A., et al., 2018. Tumor surface regularity at MR imaging predicts survival and response to surgery in patients with glioblastoma. Radiology 288, 218–225. doi:`10.1148/radiol.2018171051`.

Petersen, E., Zimine, I., Ho, Y.L., Golay, X., 2006. Non-invasive measurement of perfusion: a critical review of arterial spin labelling techniques. The British journal of radiology 79, 688–701. doi:`10.1259/bjr/67705974`.

Piechnik, S.K., Evans, J., Bary, L., Wise, R.G., Jezzard, P., 2009. Functional changes in CSF volume estimated using measurement of water T2 relaxation. Magnetic Resonance in Medicine 61, 579–586. doi:`10.1002/mrm.21897`.

Pirkl, C.M., Nunez-Gonzalez, L., Kofler, F., Endt, S., Grundl, L., Golbabaee, M., Gómez, P.A., Cencini, M., Buonincontri, G., Schulte, R.F., et al., 2021. Accelerated 3D whole-brain T1, T2, and proton density mapping: feasibility for clinical glioma MR imaging. Neuroradiology 63, 1831–1851. doi:`10.1007/s00234-021-02703-0`.

Purcell, E.M., Torrey, H.C., Pound, R.V., 1946. Resonance absorption by nuclear magnetic moments in a solid. Physical review 69, 37. doi:`10.1103/PhysRev.69.37`.

Ramalho, J., Semelka, R., Ramalho, M., Nunes, R., AlObaidy, M., Castillo, M., 2016. Gadolinium-based contrast agent accumulation and toxicity: an update. American Journal of Neuroradiology 37, 1192–1198. doi:`10.3174/ajnr.A4615`.

Ramos-Llorden, G., Arnold, J., Van Steenkiste, G., Jeurissen, B., Vanhevel, F., Van Audekerke, J., Verhoye, M., Sijbers, J., 2016. A unified maximum likelihood framework for simultaneous motion and T1 estimation in quantitative MR T1 mapping. IEEE Transactions on Medical Imaging 36, 433–446. doi:`10.1109/TMI.2016.2611653`.

Ramos-Llordén, G., Vegas-Sánchez-Ferrero, G., Björk, M., Vanhevel, F., Parizel, P.M., Estépar, R.S.J., Arnold, J., Sijbers, J., 2018. NOVIFAST: A Fast Algorithm for Accurate and Precise VFA MRI T1 Mapping. IEEE Transactions on Medical Imaging 37, 2414–2427. doi:`10.1109/TMI.2018.2833288`.

Rathore, S., Bakas, S., Pati, S., Akbari, H., Kalarot, R., Sridharan, P., Rozycki, M., Bergman, M., Tunc, B., Verma, R., et al., 2017. Brain cancer imaging phenomics toolkit (brain-CaPTk): an interactive platform for quantitative analysis of glioblastoma, Springer, Quebec, Canada. pp. 133–145.

Reeder, S.B., 2014. Gadolinium-based contrast agents: What does "single-dose" mean anymore? doi:`10.1002/jmri.24352`.

Reuter, M., Schmansky, N.J., Rosas, H.D., Fischl, B., 2012. Within-subject template estimation for unbiased longitudinal image analysis. NeuroImage 61, 1402–1418. doi:`10.1016/j.neuroimage.2012.02.084`.

Rogowska, J., Olkowska, E., Ratajczyk, W., Wolska, L., 2018. Gadolinium as a new emerging contaminant of aquatic environments. Environmental toxicology and chemistry 37, 1523–1534. doi:`10.1002/etc.4116`.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Springer International Publishing. pp. 234–241.

Runge, V.M., 2016. Safety of the gadolinium-based contrast agents for magnetic resonance imaging, focusing in part on their accumulation in the brain and especially the dentate nucleus. Investigative radiology 51, 273–279. doi:`10.1097/RLI.0000000000000273`.

Schell, M., Tursunova, I., Fabian, I., Bonekamp, D., Neuberger, U., Wick, W., Bendszus, M., Maier-Hein, K., Kickingereder, P., et al., 2019. Automated brain extraction of multi-sequence MRI using artificial neural networks, European Congress of Radiology-ECR 2019.

Segars, W.P., Sturgeon, G., Mendonca, S., Grimes, J., Tsui, B.M., 2010. 4D XCAT phantom for multimodality imaging research. Medical physics 37, 4902–4915. doi:`10.1118/1.3480985`.

Shankar, P.R., Parikh, K., Davenport, M.S., 2018. Financial implications of revised ACR guidelines for estimated glomerular filtration rate testing before contrast-enhanced MRI. Journal of the American College of Radiology 15, 250–257. doi:`10.1016/j.jacr.2017.10.008`.

Sohail, M., Riaz, M.N., Wu, J., Long, C., Li, S., 2019. Unpaired Multi-contrast MR Image Synthesis Using Generative Adversarial Networks, in: International Workshop on Simulation and Synthesis in Medical Imaging, Springer. pp. 22–31. doi:`10.1007/978-3-030-32778-1_3`.

Stanisz, G.J., Odrobina, E.E., Pun, J., Escaravage, M., Graham, S.J., Bronskill, M.J., Henkelman, R.M., 2005. T1, T2 relaxation and magnetization transfer in tissue at 3T. Magnetic Resonance in Medicine 54, 507–512. doi:`10.1002/mrm.20605`.

Stöcker, T., Vahedipour, K., Pflugfelder, D., Shah, N.J., 2010. High-performance computing MRI simulations. Magnetic resonance in medicine 64, 186–193. doi:`10.1002/mrm.22406`.

Suter, Y., Knecht, U., Wiest, R., Reyes, M., 2020. Overall Survival Prediction for Glioblastoma on Pre-treatment MRI Using Robust Radiomics and Priors, Springer. pp. 307–317.

SyntheticMR, . SyMRI IMAGE. `https://syntheticmr.com/products/symri-image/`. [Online; accessed 01-December-2023].

Tan, A.C., Ashley, D.M., López, G.Y., Malinzak, M., Friedman, H.S., Khasraw, M., 2020. Management of glioblastoma: State of the art and future directions. CA: Cancer J Clin 70, 299–312. doi:`10.3322/caac.21613`.

Tanenbaum, L.N., Tsiouris, A.J., Johnson, A.N., Naidich, T.P., DeLano, M.C., Melhem, E.R., Quarterman, P., Parameswaran, S., Shankaranarayanan, A., Goyen, M., et al., 2017. Synthetic MRI for clinical neuroimaging: results of the Magnetic Resonance Image Compilation (MAGiC) prospective, multicenter, multireader trial. American Journal of Neuroradiology 38, 1103–1110. doi:`10.3174/ajnr.A5227`.

Tewarie, I.A., Senders, J.T., Kremer, S., Devi, S., Gormley, W.B., Arnaout, O., Smith, T.R., Broekman, M.L., 2021. Survival prediction of glioblastoma patients—Are we there yet? A systematic review of prognostic modeling for glioblastoma and its clinical potential. Neurosurg Rev 44, 2047–2057. doi:`10.1007/s10143-020-01430-z`.

Thakkar, J.P., Peruzzi, P.P., Prabhu, V.C., 2021. Glioblastoma Multiforme. `https://www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Glioblastoma-Multiforme`. [Online; accessed 16-December-2021].

Thakur, S., Doshi, J., Pati, S., Rathore, S., Sako, C., Bilello, M., Ha, S.M., Shukla, G., Flanders, A., Kotrotsou, A., et al., 2020. Brain extraction on MRI scans in presence of diffuse glioma: Multi-institutional performance evaluation of deep learning methods and robust modality-agnostic training. NeuroImage 220, 117081. doi:`10.1016/j.neuroimage.2020.117081`.

Thust, S., Heiland, S., Falini, A., Jäger, H.R., Waldman, A., Sundgren, P., Godi, C., Katsaros, V., Ramos, A., Bargallo, N., et al., 2018. Glioma imaging in Europe: A survey of 220 centres and recommendations for best clinical practice. European radiology 28, 3306–3317. doi:`10.1007/s00330-018-5314-5`.

Tirosh, I., Suvà, M.L., 2020. Tackling the many facets of glioblastoma heterogeneity. Cell Stem Cell 26, 303–304. doi:`10.1016/j.stem.2020.02.005`.

Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4ITK: improved N3 bias correction. IEEE Trans Med Imaging 29, 1310–1320. doi:`10.1109/TMI.2010.2046908`.

Våge, T., 2023. Glioma Microstructure Modeling from Diffusion MRI: A Self-Supervised Deep Learning Approach. Master's thesis. The University of Bergen.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need, in: Advances in Neural Information Processing Systems.

Vasylechko, S.D., Warfield, S.K., Afacan, O., Kurugol, S., 2022. Self-supervised IVIM

DWI parameter estimation with a physics based forward model. Magnetic resonance in medicine 87, 904–914. doi:10.1002/mrm.28989.

Wang, B., Zhang, Y., Zhao, B., Zhao, P., Ge, M., Gao, M., Ding, F., Xu, S., Liu, Y., 2018. Postcontrast T1 mapping for differential diagnosis of recurrence and radionecrosis after gamma knife radiosurgery for brain metastasis. American Journal of Neuroradiology 39, 1025–1031. doi:10.3174/ajnr.A5643.

Wang, J., Chen, Y., Wu, Y., Shi, J., Gee, J., 2020. Enhanced generative adversarial network for 3D brain MRI super-resolution, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 3627–3636.

Wang, P., Hu, S., Wang, X., Ge, Y., Zhao, J., Qiao, H., Chang, J., Dou, W., Zhang, H., 2023. Synthetic MRI in differentiating benign from metastatic retropharyngeal lymph node: combination with diffusion-weighted imaging. European Radiology 33, 152–161. doi:10.1007/s00330-022-09027-4.

Wansapura, J.P., Holland, S.K., Dunn, R.S., Ball Jr, W.S., 1999. NMR relaxation times in the human brain at 3.0 Tesla. Journal of Magnetic Resonance Imaging 9, 531–538. doi:10.1002/(SICI)1522-2586(199904)9:4<531::AID-JMRI4>3.0.CO;2-L.

Wardlaw, J.M., Valdés Hernández, M.C., Muñoz-Maniega, S., 2015. What are white matter hyperintensities made of? Relevance to vascular cognitive impairment. Journal of the American Heart Association 4, e001140. doi:10.1161/JAHA.114.001140.

Warntjes, J., Leinhard, O.D., West, J., Lundberg, P., 2008. Rapid magnetic resonance quantification on the brain: optimization for clinical usage. Magnetic Resonance in Medicine 60, 320–329. doi:10.1002/mrm.21635.

Warntjes, M., Blystad, I., Tisell, A., Larsson, E.M., 2018. Synthesizing a contrast-enhancement map in patients with high-grade gliomas based on a postcontrast MR imaging quantification only. American Journal of Neuroradiology 39, 2194–2199. doi:10.3174/ajnr.A5870.

Weiskopf, N., Suckling, J., Williams, G., Correia, M.M., Inkster, B., Tait, R., Ooi, C., Bullmore, E.T., Lutti, A., 2013. Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: a multi-center validation. Frontiers in Neuroscience 7, 95. doi:10.3389/fnins.2013.00095.

Wen, P.Y., Kesari, S., 2008. Malignant gliomas in adults. New England Journal of Medicine 359, 492–507. doi:10.1056/NEJMra0708126.

Wesolowski, J.R., Kaiser, A., 2016. Alternatives to GBCA: are we there yet? Topics in Magnetic Resonance Imaging 25, 171–175. doi:10.1097/RMR.0000000000000096.

Westbrook, C., 2015. MRI at a Glance. John Wiley & Sons.

Xiao, Y., Fonov, V., Chakravarty, M.M., Beriault, S., Al Subaie, F., Sadikot, A., Pike, G.B., Bertrand, G., Collins, D.L., 2017. A dataset of multi-contrast population-averaged brain MRI atlases of a Parkinson's disease cohort. Data in Brief 12, 370–379. doi:10.1016/j.dib.2017.04.013.

Xie, H., Lei, Y., Wang, T., Roper, J., Axente, M., Bradley, J.D., Liu, T., Yang, X., 2022. Magnetic resonance imaging contrast enhancement synthesis using cascade networks with local supervision. Medical Physics doi:10.1002/mp.15578.

Yankeelov, T.E., Pickens, D.R., Price, R.R., 2011. Quantitative MRI in cancer. Taylor & Francis.

Yılmaz, E., Güldoğan, N., Ulus, S., Türk, E.B., Mısır, M.E., Arslan, A., Arıbal, M.E., 2024. Diagnostic value of synthetic diffusion-weighted imaging on breast magnetic resonance imaging assessment: comparison with conventional diffusion-weighted imaging. Diagnostic and Interventional Radiology 30, 91. doi:`10.4274/dir.2023.232466`.

You, S., Lei, B., Wang, S., Chui, C.K., Cheung, A.C., Liu, Y., Gan, M., Wu, G., Shen, Y., 2022. Fine perceptive GANs for brain MR image super-resolution in wavelet domain. IEEE transactions on neural networks and learning systems doi:`10.1109/TNNLS.2022.3153088`.

Zhang, Y., Brady, J.M., Smith, S., 2000. Hidden markov random field model for segmentation of brain MR image, International Society for Optics and Photonics, San Diego, CA, USA. pp. 3979, 1126–1137.

Zhou, J., Lal, B., Wilson, D.A., Laterra, J., Van Zijl, P.C., 2003. Amide proton transfer (APT) contrast for imaging of brain tumors. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine 50, 1120–1126. doi:`10.1002/mrm.10651`.

Zhu, A., Shih, R., Huang, R.Y., DeMarco, J.K., Bhushan, C., Morris, H.D., Kohls, G., Yeo, D.T., Marinelli, L., Mitra, J., et al., 2023. Revealing tumor microstructure with oscillating diffusion encoding MRI in pre-surgical and post-treatment glioma patients. Magnetic resonance in medicine 90, 1789–1801. doi:`10.1002/mrm.29758`.

Zhu, D.C., Penn, R.D., 2005. Full-brain T1 mapping through inversion recovery fast spin echo imaging with time-efficient slice ordering. Magnetic Resonance in Medicine 54, 725–731. doi:`10.1002/mrm.20602`.