Development of a method for MRI based prediction of microstructural features of the brain

Alhusain Abdalla
Department of Biomedical Engineering
McGill University, Montreal
June 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Engineering

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my supervisor, Prof. Amir Shmuel for his continuous support, mentorship, and valuable input to my research. I thank my parents, sisters, and brother for their emotional support throughout my masters. I also thank Prof. David Rudko for his collaboration with MRI acquisition. I thank Marius Tuznik for his diligence and professionalism during the collection of MRI data. I would like to also thank Sethu Jegathambal for his assistance throughout my masters, especially with histology acquisition and analysis. I would like to thank Roland Pilgram, who wrote the code for registration of histology images to MRI. Finally, I thank Prof. Robert Funnell, the graduate program director, for his valuable feedback and follow up.

Contributions of colleagues to this research project

The candidate served as the main researcher for this project. He carried out the review of the literature, the alignment of images of the tissue to MRI, and the deep learning analysis, compared different deep learning protocols and wrote the thesis. The following is a list of all contributors and collaborators:

- Sethu Boopathy obtained the MRI data and took part in the tissue processing.
- Dr. Aleksandra Beata Bortel performed the animal procedures for obtaining the brain tissue.
- Dr. Roland Pilgram wrote the code for registering images of histology to the MRI data.
- Marius Tuznik obtained the MRI data.
- Dr. David Rudko advised on optimization of the MRI data acquisition and fitting equations of qMRI to the MRI data.
- Dr. Amir Shmuel initiated the research project, including the use of deep learning and polynomial fitting to learn the relationship between MRI data and microstructural features of the tissue. He obtained the MRI data, advised on how to plan and perform the data analysis, evaluated and gave feedback on the results, and edited the thesis.

This research was funded by a CIHR grant and the CF-REF initiative Healthy Brain for Heathy Life awarded to Amir Shmuel. The data analysis was carried out in part using Compute Canada resources (http://www.computecanada.ca).

Abstract

This thesis presents a deep learning framework that aims to learn and predict microanatomical features from magnetic resonance imaging (MRI) of the rat brain. We obtained different MRI contrasts in an effort to understand the underlying mechanisms of MRI signals and their relationship to microstructural features. Scientists have studied numerous neurological disorders through analysing the brain's anatomy and functionality. For decades, the gold standard for accurately assessing and quantifying microstructural biomarkers of several pathologies and disorders in the brain remained to be histology. However, since acquiring histological images is only done post-mortem, this results in very limited capacity of diagnostics and prognostics. Therefore, we suggest a multimodal deep learning framework to learn and predict histological features from their corresponding MRI images. Our study forms a proof of concept and includes the preprocessing, registration, and optimization stages based on MRI data and histology that we collected in-house for this project. In addition, different analyses are conducted to test and analyze the contribution of different MRI modalities for learning specific microstructural features in the brain. We present a qualitative and quantitative performance assessment of the developed framework. Particularly, our pipeline is concerned with using T1, T2, and T2* MRI contrasts to learn and predict Nissl and Myelin histology labels through unimodal, bi-modal, and multimodal experiments. Our conclusions emphasize the significance of image resolution in such studies and shows that the performance of the models improves with an increasing number of MRI contrasts that are included in the learning.

Résumé

Cette thèse présente un système d'apprentissage en profondeur qui vise à apprendre et à prédire les caractéristiques micro-anatomiques à partir de différents contrastes d'IRM de cerveau de rat, dans le but de comprendre les mécanismes sous-jacents des signaux IRM et leur relation avec les caractéristiques microstructurales. Les scientifiques ont étudié de nombreux troubles neurologiques en analysant l'anatomie et la fonctionnalité du cerveau. Pendant des décennies, l'histologie est restée l'étalon-or pour évaluer et quantifier avec précision les biomarqueurs microstructuraux de plusieurs pathologies et troubles du cerveau. Cependant, comme l'acquisition d'images histologiques ne se fait que post-mortem, les diagnostics et les pronostics sont très limités. Par conséquent, nous proposons un système d'apprentissage en profondeur multimodal pour apprendre et prédire les caractéristiques histologiques à partir des images IRM correspondantes. Le système sert de preuve de concept et comprend les étapes adéquates de prétraitement, d'enregistrement et d'optimisation basées sur les données IRM et l'histologie que nous avons collectées en interne pour ce projet. De plus, différentes expériences sont menées pour tester et analyser la contribution des différentes modalités d'IRM à l'apprentissage de microstructures spécifiques dans le cerveau. Une évaluation qualitative et quantitative des performances du système développé est aussi présentée. En particulier, notre pipeline se concentre sur l'utilisation des contrastes IRM T1, T2 et T2* pour apprendre et prédire les étiquettes histologiques de Nissl et de myéline par le biais d'expériences unimodales, bimodales et multimodales. Nos conclusions soulignent l'importance de la résolution de l'image dans de

telles études et montrent une corrélation positive entre la performance des modèles et le nombre de contrastes IRM inclus dans l'apprentissage.

Chapter I - Introduction

Since MRI was first developed, researchers have actively pursued the imaging of anatomical structures in the brain through advancing the imaging technologies. The ability to also probe microstructural features in-vivo is anticipated to revolutionize the diagnostics and prognostics of many neurological diseases that present early microscopic markers. However, due to several hardware limitations, the ability to capture microscopic information using MRI is still very challenging. To date, histological datasets are considered the gold standard for studying the brain's morphology, microstructure, and pathology at resolutions that are not achievable by other non-invasive techniques. Histology, however, is a post-mortem procedure that cannot be deployed for living organisms. It also introduces tissue deformations and structural distortions that affect the desired *in-vivo* quantitative measurements. **Therefore, current MRI and histology imaging technologies do not provide optimal in-vivo quantitative measurements of microstructural features in the brain.**

In addition to studying microanatomical features in the brain, scientists have been intrigued to understand the underlying MR signals. Since these signals are known to be related to several physio-anatomical features, studies observing them would often need to refer to the corresponding histology data to analyze causality and correlative relationships. Overall, complementing in-vivo MRI data with ex-vivo histology data allows for better observations of microstructural features, and supports the understanding of underlying MR signals.

My thesis research is motivated by the need to detect and quantify microstructural features from multimodal in-vivo MRI data without the need to perform invasive histological procedures.

1.1 Brain Imaging

1.1.1 Magnetic Resonance Imaging

MRI is an imaging technique based on the fundamental concept of nuclear magnetic resonance (NMR). Since it was first introduced, it has been the gold standard for in-vivo imaging in several applications due to its non-invasiveness and ability to provide a wide range of anatomical and functional information through capturing multiple parameters and contrasts [1]. Many MRI studies focus on analyzing the morphometry of the brain, specifically its macroscopic and mesoscopic structures. Several longitudinal and cross-sectional studies track and measure certain morphometric properties of the cortex and sub-cortical structures, such as volume or thickness [2]–[5]. These anatomical measurements are key for studying a range of features and effects that are anatomically detectable and traceable, such as aging [6], [7], neurodegenerative disease (diagnosis and monitoring) [8], plasticity [9], and several neurophysiological endophenotypes [10]. However, several complications arise when interpreting these standard morphometry observations in the brain. For instance, the spatial resolution is often reduced due to the fact that statistical group analyses require spatial smoothing. In addition, changes observed in these studies could be due to factors other than the macroscopic variation in shape, such as image contrast influenced by microscopic processes or instrumental artefacts. These changes are often a result of a number of microstructural neuronal events, such as axonal sprouting, neurogenesis, dendritic synaptogenesis, or non-neuronal events such as vascularization [9].

The main reason behind this lack of specificity is the fact that conventional imaging does not directly map a specific MR parameter but is weighted towards it.

Weighting is a term used to describe the source of the contrast (difference in signal between different voxels) in MR images. The qualitative concept of contrast in MRI is a result of the difference in response between different tissues when specific pulse sequences are implemented. For instance, an image is considered weighted towards a tissue property such as T1 (longitudinal relaxation time), T2 (transverse relaxation time), T2* (effective transverse relaxation time), or proton density when the change of that specific property across different tissues is the key source of contrast in the image. However, the intensity values of different voxels in weighted MR images do not represent any meaningful or quantitative measurements, and are not consistent across different experiments, subjects, or locations. Therefore, conventional weighted MRI scans, such as T1-weighted, T2-weighted, or T2*-weighted contrasts may not be the best predictors or indicators of microstructures and microanatomical processes. Instead, the growing fields of quantitative MRI and in-vivo histology provide more reliable and expressive markers. Detailed review and explanation of quantitative MRI is provided in Chapter II.

1.1.2 Histology

Histology is a type of tissue imaging that studies morphology under a microscope. Since most cells are naturally transparent, histology specialists typically use a different number of stains to visualize individual cellular features and provide contrast. In brain imaging, there are a few popular stains that have been employed for varying purposes. For example, neuron cell bodies that constitutes gray matter are typically stained with Nissl. The

Nissl dye (such as Cresyl violet) attaches to nucleic acids and shows the somata with a purple color. Similarly, white matter could be studied through using a dye known as Luxol Fast Blue, that targets and stains myelin – the protective fatty layer that surrounds neuronal axons. Another staining technique is immunohistochemistry, which relies on the fact that antibodies are specific to certain antigens (molecular targets) found on certain cell types. It thus uses antibodies fused to fluorescent compounds or enzymes to highlight and visualize target sites using fluorescence or enzymatic reactions.

A number of histological procedures are implemented before staining the tissue to ensure optimal results and visualization. First, the brain is perfused to preserve its cellular structure. Perfusion refers to draining the blood from an animal's vasculature and using it to pump a chemical fixative into the tissue. The brain gets fully soaked in fixative after dissection to ensure preservation process is complete and begin the embedding process. The embedding process involves inserting the brain in a medium – such as paraffin - that possesses similar mechanical properties to the tissue. Following that, a sectioning instrument called microtome is used to cut the brain into thin slices. These slices are then left to dry after mounting them on transparent slides. After preparing and slicing the tissue, the staining procedure takes place, and the slices are then imaged using appropriate microscopes.

1.2 Rationale behind complementing MRI with histology

Structural MRI has the potential to non-invasively estimate the brain's microanatomical features. With the rapid advancement in medical imaging technologies, it is now possible to image the brain in-vivo at a single-cell level in animal models using 7 Tesla

MRI [11]. However, up until this day, microscopic resolutions at the level of neurons or synapses are not yet possible to directly image using in-vivo MRI [12], [13]. It is indisputable that post-mortem microscopy data and histology are the gold-standard in providing the maximum amount of microstructural information at the highest resolutions. Therefore, complementing histological information with non-invasive in-vivo imaging has been a highly pursued task over the years.

Studies that looked at the diagnostics and prognostics of cancer pathogenesis has driven the advancements in several in-vivo imaging techniques, such as computed tomography (CT), single-photon emission computed tomography (SPECT), positron emission tomography (PET), and MRI. These techniques are able to non-invasively and quantitatively identify and measure a range of molecular, functional, and structural features of oncological lesions [14], [15]. Of these methods, MRI stands out by offering the maximum possible amount of information simultaneously from multimodal functional and morphological contrasts, which allows simultaneous monitoring of carcinogenesis and the different processes corresponding to its pathway [16]. For instance, prostate cancer imaging has been facilitated significantly through multiparametric MRI. Specifically, higher spatial resolutions (<1mm) and the comprehensiveness of information provided from both morphological and functional images allowed prostate cancer MR examinations to characterize vascular and cellular features of possible lesions [17].

Several limitations arise from the utilization of MRI to assess oncological lesions, which clearly reflect on the diagnostic predictive power, treatment response prediction, diagnostic accuracy, and differential diagnosis [18], [19]. Despite being invasive, histology is

still considered the gold standard in characterizing and assessing lesions, and the absence of histological information that corresponds to MR images is the main source of the afore mentioned limitations. When analyzing correlations between histological data and imaging, the bias introduced from four main sources need to be addressed [20]. First, when a histological sample is taken outside of a living organism, tissue deformations occur due to the absence of mechanical compression and tension that existed inside the organism from all the surrounding structures and vascularization. This necessarily imposes local and non-linear inconsistencies in the alignment between histology and MRI images. Second, a significant mismatch could be introduced due to different orientations between the histological sample and the imaging planes of the scans. Third, since the two methods use different spatial resolutions (MRI: 0.5 - 3 mm, histology: 1 - 5μm), a perfect superposition of the images cannot be reached. Finally, it is difficult to assess the registration accuracy between the two imaging modalities due to their different contrast resolution.

The ability to extract histological information from in-vivo MRI scans will open new possibilities and help overcome several limitations in the current standard of imaging. To start with, the in-vivo non-invasive assessment of microstructural features will advance the diagnosis and prognosis of several brain disorders and pathologies. Moreover, collecting the measurements in-vivo will reflect the real-time effects and changes in microstructural features, and would eliminate distortions related to extracting histological samples post-mortem outside of its anatomical environment. Finally, collecting structural MRI and micro-structural information in single imaging session is far more efficient in terms of cost, time, and convenience.

1.3 Objectives

The key objectives of my work are to:

1) Study the relationship between the MR signal and the corresponding microanatomical features presented by immunohistochemistry.
2) Develop a deep learning framework that predicts microanatomical features of the brain from MRI images without having to resort to invasive or post-mortem approaches.
3) Test and validate the deep learning solution using different datasets and through performing different experiments. These experiments will include uni-modal, bi-modal, and multi-modal MRI input to predict individual microanatomical features.

1.4 Proposed Solution and Approach

The ability to infer microanatomical features of the brain from MRI scans will open numerous possibilities in the field of neurological diagnoses as well as help elucidate the complex relationship between the macroscopic MR signal and the underlying microanatomy. To do that, we propose a deep-learning-based framework to predict microanatomical features from quantitative MRI without the need for the corresponding histology data. However, constructing an accurate predictive model for multimodal data is not a trivial task. We started by collecting in-vivo and ex-vivo MRI data for T1, T2, and T2* images from a rat. On the same animal, we performed histology staining and imaging for myelin, iron, and Nissl. Before creating the deep-learning model, we had to extract 2D MRI images from the 3D MRI volumes and register them with their corresponding histology slices to achieve a reliable alignment. This allowed us to start training the deep-learning network and optimizing its

hyperparameters to obtain a highly representative model of the data. We trained various models using different combinations of the data to determine the effect of individual MRI parameters or histological stains on the quality of the prediction. Finally, based on the results of these different models, we provide our analyses and conclude with possibilities for future work.

Chapter II- Review and Background

2.1 Diseases identified by imaging

Several pathologies and neurodegenerative diseases are identified and monitored using MRI, such as lesions, multiple sclerosis (MS), and epilepsy. MS is an immune-mediated disease in which the central nervous system (CNS) is attacked by the immune system of the body itself [21]. MS could be identified by demyelinated inflammations that present as lesions in the brain. Initial diagnosis and monitoring of MS are commonly performed using MRI. However, the CNS lesions presented as MRI abnormalities are not specific for a certain pathology but could be originate from a multitude of biological processes, such as demyelination, gliosis, axonal loss, or inflammation. Currently, globally approved treatments of MS focus on slowing the rate of myelin damage through suppressing the inflammation. However, these treatments do not cure MS nor stop its progression and are only anti-inflammatory. Recently, treatments that target myelin reparation instead of inflammation suppression are being studied and tested. Myelin is a protective layer made of protein and fatty substances that insulates axons to enhance their conductance. Damage in myelination hinders the flow of electrical impulses through neurons, which can lead to axonal dysfunction and several neurological symptoms [22]. Assessment of new myelin treatments and their efficacies requires accurate and quantitative in-vivo measurement of myelin. In-vivo MRI-based quantification of myelin has been the focus of many studies as described in the review by Heath et. al [23]. Although these studies and MRI methods have advanced significantly in myelin quantification, it is still necessary to compare and evaluate them against the current gold standard, i.e., histology. This comparison will give insight into the accuracy, specificity, and reproducibility of the techniques. A systematic review performed

by Weijdan et. al [24] on MRI-based myelin quantification techniques concluded that they generally correlate well with corresponding histology data. In humans, the authors concluded that the myelin-water fraction (MWF) technique correlated the best with histology (R^2 = 0.85). However, it was also concluded that the accuracy and resolution of MRI-based estimation of myelin was still not close to that of the histology. In addition, there was insufficient data provided by MRI-based myelin quantification studies and the performance variability was too high across studies, not allowing the review authors to make a very reliable conclusion about those methods [24].

Another common neurological condition that could be detected and monitored through MRI is epilepsy. Epilepsy is a disorder in the CNS caused by abnormal brain circuitry and leads to seizers and uncontrollable behavior for short periods of time. Statistics showed that epileptogenic lesions that were identified using MRI had more favorable surgical outcomes. [25], [26]. However, studies showed the sensitivity of preoperative conventional MRI scanning protocols was unreliable in detecting seizure onset lesions, where more than 30% of patient diagnosis results were false negative [27], [28]. Quantitative MRI mapping, such as T2 and diffusion tensor imaging (DTI) [29]–[31], showed better correlations with the histology when compared to conventional weighted MRI.

Analyzing the correlations between MRI and histopathology of the neocortex in temporal lobe epilepsy (TLE) patients helped build better conclusions about the imaging-histopathological relationship. A study conducted by Garbelli et. al [32] correlated a visible blur in the temporal pole cortical boundaries on fast fluid-attenuated inversion recovery (FLAIR) and T2-weighted images with the degeneration of fiber bundles. Another study by

Eriksson et al [33] that the gray matter (GM) on FLAIR and T2-weighted images was negatively correlated with neuronal nuclear antigen (NeuN). However, a follow-up study found no correlations between probability maps of GM and NeuN or glial fibrillary acidic protein (GFAP) [34]. Another follow-up study that looked at DTI and FLAIR also found no correlations [35]. This discrepancy in results and data suggests a poor understanding of the abnormal MR signals from focal epilepsy in light of their pathological basis. Quantitative imaging show promise in understanding the relationship between seizure outcomes and specific MRI parameters in patients that were classified as false negative using conventional weighted MRI [36]. In-vivo identification of these histopathological microstructures should also improve the detection, monitoring, and surgical outcomes of hippocampal sclerosis patients.

2.2 Brain microstructure and MRI parameters

In-vivo visualization of brain structures have recently improved significantly due to advancements in MR imaging, where it is now possible to image human brains at 350 μm isotropic voxels within feasible scan durations. Since the contrast of MRI images originates from a number of different microanatomical structures and processes, it is important to understand them and quantify their effects on MRI observations. Figure 1 demonstrates an example of the multitude of features and structures held within a single brain slice and affects the MRI contrast at different levels. Myelin and iron are considered the top contributors to MRI contrast and will be highlighted more in this review.

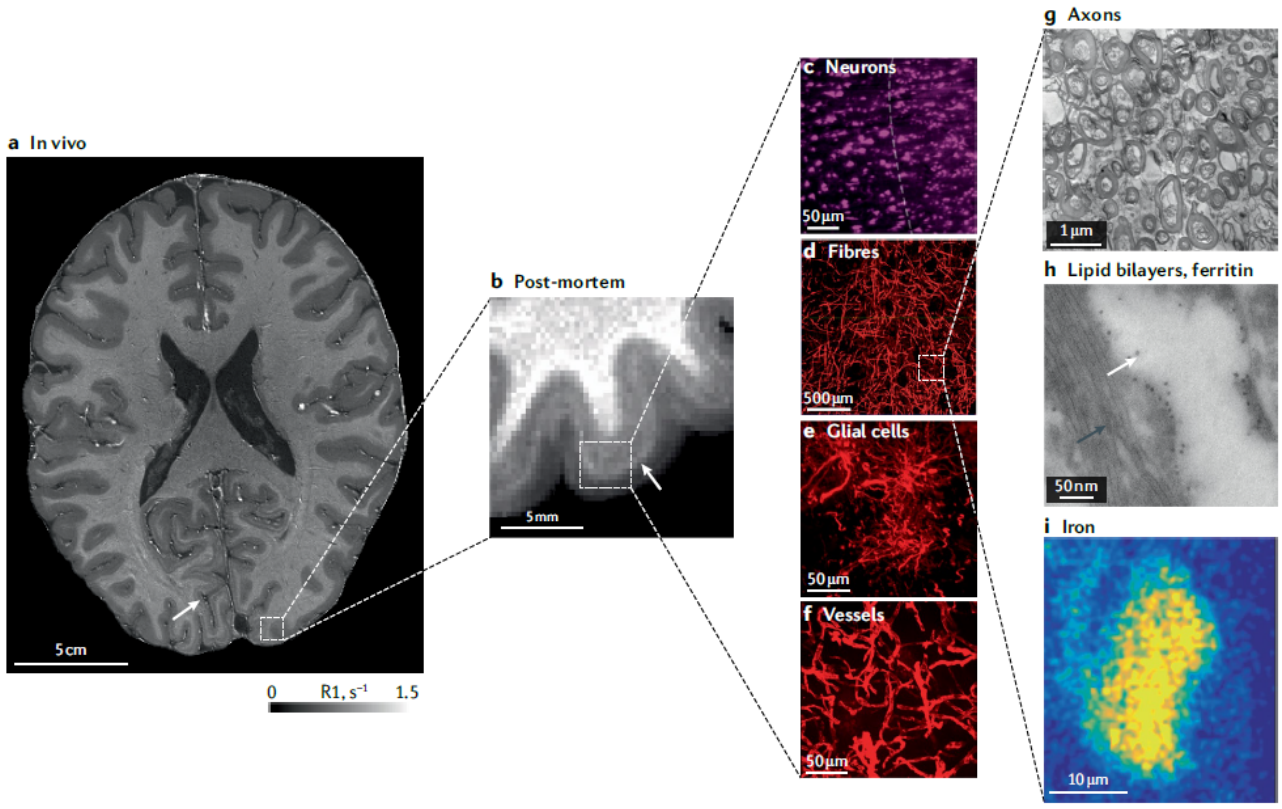

Figure 1: Structural information extracted from a single brain slice at multiple scales. (Figure reproduced from Weiskopf et al. [1] with permission)

Myelin comprises approximately of 70% lipids (30% phospholipids, 20% galactolipids, and 20% cholesterol) in addition to at least 340 different kinds of proteins that make up the remaining 30% of myelin's dry weight [37]. The relationship between myelin and the MR signal has been investigated by Laule et. al. [38] through measuring the MWF in MRI images of MS patients and using cadaver brain slices stained for myelin. Similarly, iron and its distribution in the human brain have been investigated using standard staining procedures, such as Turnbull's blue for divalent iron and Perl's Prussian blue for trivalent iron [39]. However, these techniques are not very reliable and are known to present some biased distribution of iron. As for myelin, the quality of staining is also known to vary

significantly based on several factors. For instance, the post-mortem duration before fixating, the type of fixation, the stain type, and the stain bath temperature all play important roles in preserving the quality and reliability of the staining.

Several studies have reported a relationship between myelin content and the T1, some of which only offered vague statements and observations with no supporting quantitative evidence [40], [41], while others attempted to find a quantitative relationship between myelin and T1 [38]. Another study by Glasser et al. [42] claimed that the ratio between T1w and T2w images can be used to map myelin. However, since current imaging technologies are not yet able to determine an absolute measurement of myelin concentrations, quantifying such relationships is still an active area of study. Similarly, the contribution of iron in the formation of T1 contrast is still ambiguous [43]. The cortex, for instance, often has concentrations of both myelin and iron, and the T1 contrast there is not only caused by myelin but also partly by ferritin molecules (iron) according to a study [44]. However, studies on the contribution of iron in T1 contrast contradict one another in many cases. Some are reporting no significant effect of iron on T1 [45], while others present data that shows a correlation between iron and T1 [46], [47].

T2* relaxation describes the effective decay in transverse magnetization caused by both the spin-spin relaxation and magnetic field inhomogeneity [48]. Water on T2w images appear hypointense when it is close to paramagnetic iron due to its local magnetic field gradients, especially in areas with higher iron concentrations [49]. Research focused more on the relationship between iron and T2* than any other MR parameter because of its relevance in clinical applications, such as visualizing iron for Parkinson's disease [49]. It has

been suggested by earlier studies to use T2*w brain images as biomarkers for iron mapping [50]. However, recently, quantitative susceptibility mapping (QSM) has been getting more focus as it can better visualize the distribution of iron in the brain tissue [51].

2.3 Quantitative MRI

QMRI is an imaging technique used to quantify specific MR tissue properties, such as magnetic susceptibility [52]–[55], relaxation times [43], [56]–[59], proton density [60]–[62], and tissue exchange processes [63]–[66]. These physical parameters are mainly determined by macromolecular concentrations. Myelin, as discussed earlier, is an encasing sheath that protects axons which has high macromolecular content. This means that several parameters such as magnetization transfer (MT) and relaxation times (longitudinal T1 and transverse T2) [62], [67], [68] are markers for myelination in the brain.

Diffusion imaging [69], [70] captures quantitative information of microscopic structures that hinders or restricts water diffusion, such as axonal myelin sheaths [71]. This allows quantitative measurements of microscopic features such as neurite densities and axonal diameter [72], [73], as well as fibre tracts delineation [74]–[77]. MT imaging also captures information regarding microanatomical features such as myelin through observing its effect on mobile protons visible on an MRI scan [78].

Biases that exist in conventional structural MRI scans such as radiofrequency (RF) coils' transmit and receive profiles [61], [79], [80] are eliminated from qMRI. In addition, contrast-generating parameters are quantified separately, unlike standard contrasts that are a mix of several MR parameters, such as MT ratio, T1-weighted, or T2-weighted images [81], [82]. Therefore, microstructural features are generally better estimated in qMRI.

Morphometric analysis and segmentation of particular subcortical regions on T1-weighted scans is often a challenging task. Quantitative MT maps have been utilized to overcome some of these challenges through its enhanced specificity [81]. Voxel-based morphometry (VBM) effectively translated the contrast between the white matter, substantia nigra, red nucleus, and basal ganglia into a higher sensitivity. High resolution maps of MT (800mm isotropic) and proton density (PD) facilitated multiparametric segmentation that improved the delineation process and even allowed for the brainstem to be effectively parcellated into separate classes of tissue [83]. Subtle variations in histological features that were captured on qMRI yet concealed in standard structural MRI illustrate the sensitivity of qMRI. This increased sensitivity is validated by the clear correspondences between ground truth brain anatomy and the aforementioned morphometric results. Several studies showed that the interpretation of morphometric measurements could be refined using qMRI. For instance, Lorio et al. [84] concluded that the age-related volume reduction in the gray matter of the basal ganglia observed using VBM is consistent with the underlying histological changes and the corresponding T1-weighted change in contrast.

Conventional MRI methods provide poor comparability and reproducibility across time points and sites, which significantly affect longitudinal and multicenter studies [80]. On the other hand, qMRI offers a high degree of reproducibility and comparability through standardized voxel values across different times and sites. A study demonstrated this highly equivalent comparability through imaging for quantitative T1 and T2 maps at 1.5 Tesla across different MRI scanners from different manufactures and at different imaging centers [85]. Another study mapped for T1, T2, PD, and MT parameters at 3T across three imaging sites and the results were highly aligned [61]. Unlike conventional T1-weighted intensity

values, qMRI provides microstructural biomarkers that allow reproducible comparisons of data points across time and sites as shown in Figure 2. Therefore, data for group studies across different scanners that is by default standardized in quantitative measurements could be pooled and used to investigate microstructural variations related to ageing, for example. In fact, reproducibility and comparability could even be achieved across field strengths through the use of PD maps or different derived measurements such as, for example, macromolecular tissue volume [86].

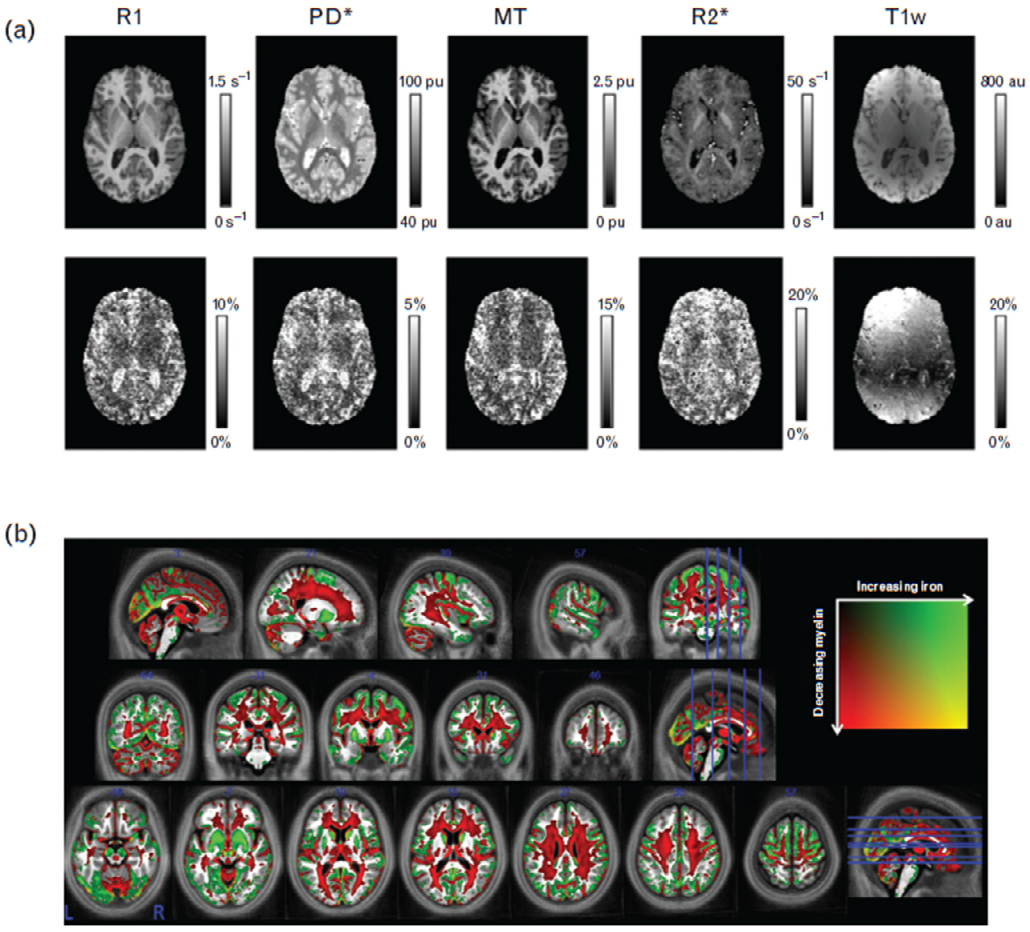

Figure 2: Applications of qMRI. (a) The mean and coefficient of variation (CoV) of inter-site qMRI maps and T1w images. (b) Myelin and iron concentrations differences related to age and estimated from qMRI biomarkers of T1, T2, and MT maps. (Figure reproduced from Weiskopf et al. [87] *with permission).*

With the increasing interest and advancements in qMRI, a wide range of opensource toolboxes have emerged to facilitate the processing and the different applications of qMRI studies. Even though most toolboxes are focused on the analysis of the brain, they are not only limited to that. hMRI is a popular toolbox developed in 2019 and is based on MATLAB and the Statistical Parametric Mapping (SPM) package [88], [89]. It enables accurate estimation of key qMRI maps, such as MT saturation, PD, T1, and T2*. The other MATLAB-based toolbox mRQ [86] on the other hand, allows the processing of the same qMRI maps in addition to the quantification of macromolecular volume, water-surface interaction rate, and interacting water protons apparent volume. Finally, a comprehensive toolbox developed by Karakuzu et al. [90] using MATLAB allows the processing of all relaxometry maps, quantitative susceptibility maps, quantitative MT, diffusion imaging, and field mapping (B0 and B1+). It also offers a user-friendly graphical user interface (GUI) for visualization and simulation as well protocol optimization routines.

2.4 In-vivo histology (hMRI)

In-vivo characterization of microstructural parameters in the brain (e.g. axonal diameters) has further improved by advancements in biophysical modelling and interpretation of the MR signal. Since these measurements are typically only possible to obtain through ex-vivo histology, researchers studying this developing method are calling it in-vivo histology or hMRI. The main challenge that hMRI is aiming to tackle is the ability to directly estimate the microstructural features and histological markers of brain tissue, which is now more achievable due to recent advancements in biophysical models and MRI acquisition. For instance, higher spatial resolutions and better data quality are direct results of improved gradient strength and RF coils [91], higher field strengths [92], optical

prospective correction of motion [93], and retrospective correction of instrumental, physiological, and subject artefacts [94], [95]. In addition, the effective resolution could be further increased through super-resolution methods [96], [97] and adaptive de-noising [98]. MRI and qMRI measures could be transformed into in-vivo histological maps via biophysical modelling as shown in Figure 3. This could be accomplished by building direct relationships and inferring upon microscopic biological features that typically require ex-vivo histological processes or analysis to obtain, such as measuring the axonal g-ratio or staining for myelin concentrations.

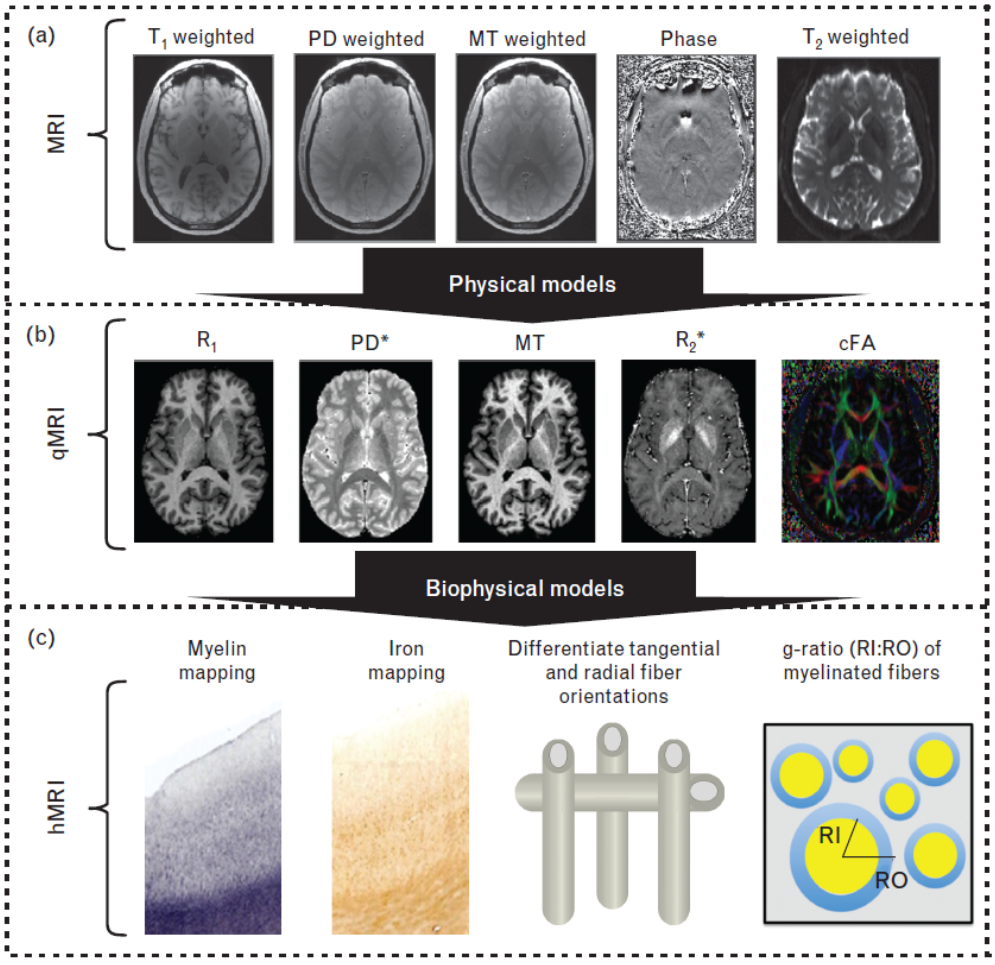

Figure 3: Using MRI and qMRI measures to obtain in-vivo histological measures. (Figure reproduced from Weiskopf et al. [87] *with permission)*

A correlation between one physical MR parameter in brain tissue and the concentration of myelin or iron is the basis of the simplest biophysical models. For instance, an increase in the concentration of myelin in a tissue is directly correlated to an increase in MT and to shorter T1, T2, and T2* relaxation times [62], [67]. On the other hand, an increase in the concentration of iron in a tissue corresponded to a decrease in T1, T2, and T2* relaxation times [43], [58]. Diffusion weighted imaging (DWI) shows that axonal degeneration and demyelination directly results in reduced values of fractional anisotropy [99], [100]. It is noted that these biophysical models describe the effect that is visible on MR signal and parameters caused by the underlying microstructural processes. However, the reverse relationship of inferring microstructural features from MRI could be misleading and complicated to interpret. For instance, an increase in myelination might be mistaken by an increase in iron concentration on T1 maps, since they both result in a shorter T1. Also, blood depositions and calcification may be mistaken for one another on T2-weighted scans since they both appear identical [101].

Since several microstructural variations could affect qMRI measures, it is necessary to combine multiple qMRI measures to be able to extract meaningful and quantitative microstructural or histological features of a brain tissue. In order to model this multiparametric relationship, several studies have investigated multiple regression and multivariate models. A study by Callaghan et al. [102] used a linear model to describe the relationship between MT, T1, and T2 maps, which allowed the inference of one parameter from the other two. For instance, MT and T2 could be used to estimate T1, which allows the synthetization of additional qMR maps to improve the correction of artefacts or reduce computational time.

Stuber et al. [103] studied the relationship between the concentration of histological features (myelin and iron) and MRI parameters (T1 and T2) using a linear multivariate model. Direct estimation of the content of myelin, iron, and potentially other histological measures from MRI parameters could be feasible through this linear model, which calculated hMRI measures comparable to the measures calculated with proton-induced X-ray emission (PIXE). However, this might only be applicable in the brain regions included in the study (the subthalamic nucleus, the primary visual cortex – V1, and the pre/post-central gyrus – M1, S1). These models developed by Stuber [103] and Callaghan [102] represent a solid step towards non-invasive direct in-vivo quantification of microstructural characteristics and measurements in the brain.

More complex estimations of specific sub-voxel properties derived from the MR signal require more complex qMRI models. For instance, magnetization transfer effects are modelled using multi-compartment models ('two-pool models') [65], [79], [104]. The same type of modelling is also used to describe axonal features from DWI that are key factors in the velocity of nerve conductance, such as axonal diameter and fiber density [71]–[73], [105]–[107]. The recent improvements in acquisition times [108], [109] and developments in high-power gradient technology [91] can facilitate the increased need of data for these models.

Susceptibility weighted imaging is an MRI contrast that has gained popularity over the past few years [54], [110] due to its ability to directly estimate non-haem iron, which is a marker for myelin breakdown and neurodegenerative processes, and is important for maintaining myelination patterns and brain function [111]. This contrast, which is obtained

from the often-overlooked intrinsic phase data of the MR signal, is unique for its potential to characterize entities that are not possible to characterize elsewhere on other contrasts [101]. Magnetic susceptibility can be quantified from the same data used to extract quantitative T1 and T2 maps. Wharton and Bowtell [53] studied magnetic susceptibility and provided a tensor description that is analogous to how water diffusion was modelled using mutli-compartment models, which opens the door for a combined modelling of contrast mechanisms. Although susceptibility imaging holds great potential for characterizing brain tissue and its microscopic features with MRI, its underlying mechanisms is not fully understood and is still an active area of research [112]–[114].

2.5 Registration of MRI and histology

Image registration generally refers to the spatial alignment of two images with similar features or structures in the same frame. Those two images usually differ in one or few characteristics, such as shapes within the image, intensities, viewpoint, deformations of structures, or noise. Numerous methods have been developed to tackle down this problem, starting from manual or semi-automated implementations all the way to fully automated pipelines [115].

Registration algorithms estimate a transformation that maps voxels from one image to another such that most corresponding voxels align if both images were superimposed on top of each other. This registration process involves three main components: a geometric mapping that describes the spatial transformation between the voxels, a metric that evaluates the quality of the registration, and an optimization algorithm to tune the mapping parameters until the best combination of parameters is found [116].

Transformations that used to map images in a registration problem can be classified into two types: rigid and non-rigid transformations. Rigid mappings maintain parallel planes and preserves a relative angle between the lines. These transformations are calculated using parameters with 6 degrees of freedom (DOF), controlling the translations and rotations (pitch, roll, and yaw) along three orthogonal axes [117]. According to literature, rigid transformations are typically used in two scenarios: when registering rigid structures such as bones [118], or as a pre-registration step before applying a more complex transformation [119]. Non-rigid transformations on the other hand include similarity transforms (allows for additional uniform scaling), affine transformations (allows for additional shear), curved, and projective [115]. There are several methods used to interpolate intensity values to new voxels, such as nearest neighbor, trilinear, and sinc interpolations [117]. There is also a range of optimization functions that are used to find a spatial mapping that minimize the loss function or the quantitative dissimilarity between the two images. Common functions include correlation, mutual information (MI), mean squared error (MSE), and other variations of these metrics.

The registration of medical images from different modalities is usually governed by a few additional factors, such as the differences in contrast and resolution. For instance, registering histological images to in-vivo MRI images might be complicated by the fact that in-vivo MRI typically has a much lower resolution than histology, which makes it difficult to trace corresponding features across the two images. A solution to this that has been described in many publications suggests using an intermediate modality to bridge the differences between the two targeted modalities [117]. In this case, ex-vivo MRI is used as the intermediate image since it has the same contrast and was acquired using the same

technology as in-vivo MRI but has higher resolution and can better correspond to features and structures on the histological image. Another complication that usually arises from this registration setting is the difference in the acquisition imaging planes [115]. Addressing this requires ensuring that the sectioning of the brain is done in the same orientation of the MRI. Using a device to help fixate the brain in a specific orientation and obtain slices that are evenly distributed has been proposed in previous implementations [117]. Other studies have suggested 3D-printing an MRI-based customized concave model to hold the brain during sectioning [117], [120]. This ensures that the sectioned slices are already coarsely aligned to the MRI data.

2.6 Image-to-image translation

The use of artificial intelligence tools for pattern recognition and feature extraction applications have amplified significantly across all fields. Unlike conventional statistics, machine learning models focus on predictions rather than inferring relationships between variables or features and are particularly powerful for large datasets complicated non-linear relationships. However, machine learning requires structured data with pre-engineered features to perform optimally. This is not the case for images, for instance, where no "structured" sense could be directly inferred of how pixels are organized in rows and columns. Therefore, deep learning and neural networks was introduced as a sub-category of machine learning to offer a sophisticated yet efficient approach that mimics the complicated connections of neurons in the brain to automatically engineer features and build non-linear models from non-structured data. Even though these neural networks generally require vast computational power and more data to perform optimally, they remain the gold-standard today in image-based learning and analysis.

There are usually two main prediction tasks that are performed by any machine learning or deep learning model: classification or regression. Classification refers to the labeling of input data into categorial classes or discrete quantities, such as predicting the existence of a malignancy or an abnormality from an input MRI scan. Regression on the other hand predicts a continuous output or quantity, such predicting the age of a patient based on their MRI scan. Image-to-image translation is sub-application of image regression, which maps and predicts continuous values of pixels from one image domain to another image domain rather than one single output value. This concept has been key in several implementations across different fields such as segmentation, image synthesis, pose estimation, restoration, and style transfer.

With the rapid advancements in computational analyses, interest in cross-domain image synthesis in the field of medical imaging has increased. Image synthesis, transformation, or mapping refers to the accurate procedure of estimating the corresponding image b2 of a subject in D2 (target domain) from an image b1 of the same subject in D1 (source domain). There are generally two main tactics to go around this problem: intensity-based methods and registration-based methods [121].

Methods based on registration begin by registering the images to create atlas of co-registered images a1 and a2, acquired in domains D1 and D2 respectively [122]. The underlying assumption behind these methods is that a geometric wrap relates images of different subjects within the same domain. For instance, if the task is to synthesize b2 from b1, an estimation of the wrap that maps a1 to b1 is applied to a2. However, registration-based methods are not effective when the underlying morphology across different subjects

is very different, since the methods are solely based on geometric transformations [121]. For example, a failure could occur due to an inconsistent pathology across the atlas or across test subjects. Also, even within normal subjects, the accuracy of registration in a single domain can be limited.

Intensity-based techniques are powerful alternatives that are not solely based on linking the anatomies of different subjects by geometry [121], [123]–[126]. For instance, frameworks of compressed sensing provide an effective approach for synthesising MRI images from different contrasts, where the patches of an atlas image a1 is linearly combined in a sparse representation and are used to express every single patch in b1, the source image [125]. The patches of image b2 are then estimated from patches of image a2 through applying these learned sparse representations. Generative models based on labelled tissue segmentation and multi-scale patches were also used to improve cross-domain patch matching [127], [128]. Other studies demonstrated that predicting image b2 from image b1 can be possible through learning generalized mappings based on nonlinear representations of voxels in image a2 based on patches in image a1. Techniques such neural networks [129] and nonlinear regression [121], [123], [124] could be utilized to learn this nonlinear mapping from training data. For instance, a supervised machine learning method called REPLICA has been developed by Jog et. al [121] to perform image synthesis based on random forests and multiresolution patches. This technique shows promise in translating MR images across contrasts or across different modalities. However, since the dictionary constructed deals with patches at different resolutions independently and the algorithm averages the separate predictions from different forest trees, the performance of the synthesis is suboptimal and high-frequency spatial information is lost.

Recently, an end-to-end deep learning approach called Multimodal was proposed for image synthesis of MRI contrasts [130]. This framework predicts a target image contrast based on several source images contrasts using a trained neural network. Moreover, the Multimodal framework constructs a dictionary of multiresolution patches and performs image synthesis under a single umbrella, which demonstrated satisfactory results – even with limited datasets – compared to other conventional techniques. However, the Multimodal method requires the input MRI contrasts to be perfectly registered spatially for optimal performance. Also, high spatial-frequency errors are not effectively captured by the absolute or mean squared error loss functions. [131]–[133].

Chapter III - Development Methodology

3.1 Pipeline

As described earlier, our pipeline consists of four main components: Data collection and processing, registration, model generation, and model evaluation. Figure 4 explains the different stages of the pipeline.

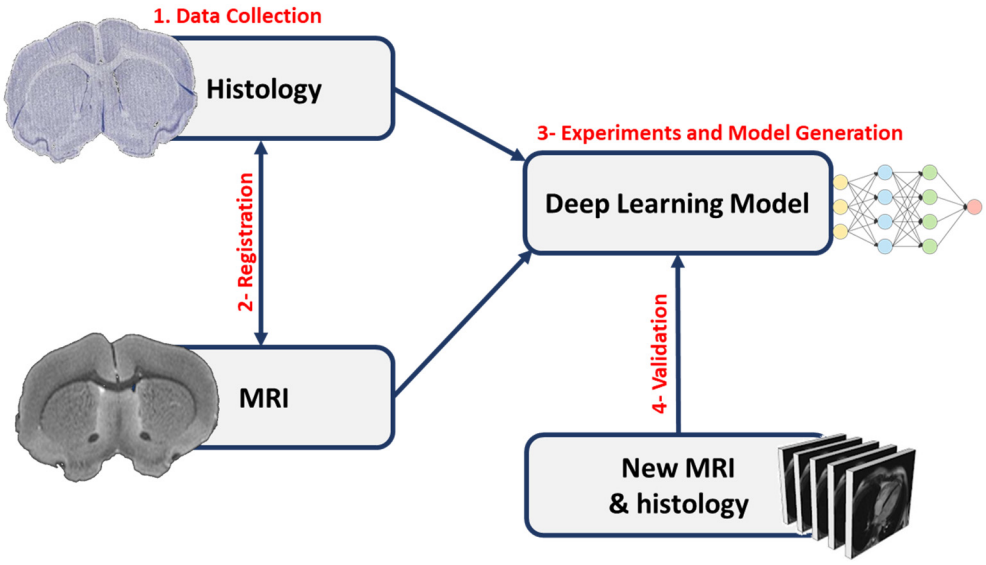

Figure 4: Summary of the pipeline of our framework

1- **Data Collection and processing:** In this stage, we collected in-vivo and ex-vivo MRI data and the corresponding histology data from the same brain to analyze the relationship between two different microstructural labels of interest – Nissl and Myelin (we also stained for iron – but preferred to not include the results, as will be explained later in this section). We also preprocessed the data to prepare it for registration and deep learning. This included skull stripping, noise reduction,

intensity normalization, and other morphological filters to ensure emphasis on key features and preservation of information necessary of the following steps.

2- **Registration:** In this stage, we aligned the different datasets in the same space using image registration techniques and a tailored pipeline to ensure optimal feature mapping between MRI and histology. We validated the registration results using both quantitative and qualitative evaluation methods.

3- **Experiments and Model Generation:** We developed and optimized a GAN-based deep learning algorithm to accurately map histological features to MRI parameters. We also designed several experiments for analyzing the contribution of histological features to shaping the MR signal of different individual contrasts. The experiments include unimodal, bi-modal, and multimodal training of models using different sets of data.

4- **Validation:** Finally, after the different models are generated, we used a randomly selected testing dataset from images that were never seen by the training algorithm and that cover different areas of the brain to evaluate the performance and prediction quality of these learned models. To visualize and quantify the performance, we used a set of qualitative and quantitative evaluation techniques. We then summarized the results and determined the best predictive MRI parameters of microstructural features.

3.2 Data

3.2.1 MRI

To the best of our knowledge, there are currently no publicly available datasets of rat brain MRI and histology from the same animal. Since morphological and microanatomical

features vary due to sex, age, individual biological differences and other factors, it was not optimal to rely on the publicly available datasets alone to accurately test our framework and map the features across MRI and histology. Therefore, we collected our own complete dataset from a 3 months old female Sprague Dawley rat, including in-vivo and ex-vivo MRI of multiple contrasts, and histology of multiple stains.

Since our pipeline depends on mapping features from one modality to another, we were opting for the best MRI resolution and contrast we could obtain given our available resources. After a few rounds of trials and improvements, we finalized the acquisition protocols and parameters that yielded satisfactory scans. Both in-vivo and ex-vivo acquisitions were performed on a 7 Tesla Bruker Pharma-scan MRI system. We acquired multiple scans using Mult-echo Spin-echo (ME-SE), Multi-echo Gradient-echo (ME-GRE), and Variable-TR (VTR) protocols to obtain T2, T2*, and T1 images respectively. We used the FieldMap method on ParaVision 5.1 (i.e. the imaging software used to acquire data on the Pharmascan system) to acquire B0 maps and to shim B0 at the outset of the experiment. We also acquired B1+ maps using 2 spin-echo echo planar imaging (SE-EPI) sequences with 2 different flip angles for the excitation pulse. For in-vivo MRI, the rat was under anesthesia, laid in a 'sphinx'-like position in a custom-made MRI bed for rats. The temperature of the animal was monitored and kept at approximately 37 degrees using warm air flow into the rat bed. The spontaneous respiration of the animal was monitored by means of a small air cushion positioned below the chest and connected to a sensor. For transmission, we used a Bruker 1 Tx channel, 70 mm ID coil. For receive, we used a Bruker 1 Rx channel, 38 mm ID coil. *Table 1* describes the acquisition parameters associated with each of the in-vivo scans. For ex-vivo MRI, we used a smaller coil of 2.3 cm diameter (we could use this small coil

because we scanned only the brain, after extracting the brain from the skull and scalps) and acquired 9 repetitions per each protocol. We later averaged these 9 repetitions for better Signal to Noise Ratio (SNR). *Table 2* presents the acquisition parameters of an individual cycle in the ex-vivo scanning protocols. For ME-GRE scans, there was an interpolation in slice direction, which explains the mismatch between the slice thickness and the 3rd dimension of the voxel size.

Table 1: Summary of imaging parameters for in-vivo MRI acquisitions

Method	Target	No. of Averages	Voxel size (*um*)	Slice thickness (*um*)	TR (*ms*)	TE (*ms*)	Flip Angle (*degree*)	FOV *(mm)*
ME-GRE (3D)	T2*	2	200x400x200	400	40	[1.5647, 4.5647, 7.5647, 10.5647, 13.5647, 16.5647, 19.5647, 22.5647, 25.5647, 28.5647]	22	[3.2, 4.8, 3.2]
MSME (2D)	T2	4	200x200	400	2338.096	[9.6365, 19.273, 28.9095, 38.546, 48.1825, 57.819]	180	[5.01, 3.5]
RARE-VTR (2D)	T1	1	200x200	400	[8000, 6000, 4000, 3000, 2000, 1200, 800, 400,]	9	180	[5.01, 3.5]

Table 2: Summary of imaging parameters for ex-vivo MRI acquisitions

Method	Target	No. of Averages	Voxel size (*um*)	Slice thickness (*um*)	TR (*ms*)	TE (*ms*)	Flip Angle (*degree*)	FOV *(mm)*
ME-GRE (3D)	T2*	4	100x300x100	300	40	[1.5647, 4.5647, 7.5647, 10.5647, 13.5647, 16.5647, 19.5647, 22.5647, 25.5647, 28.5647]	22	[3.24, 1.5, 1.6]
MSME (2D)	T2	4	100x100	300	5200	[9.6365, 19.273, 28.9095, 38.546, 48.1825, 57.819]	180	[3.2, 1.5]
RARE-VTR (2D)	T1	1	100x100	300	[8000, 6000, 4000, 3000, 2000, 1200]	9	180	[3.2 1.5]

Even though we collected the protocols necessary to extract quantitative MRI data, we decided to save it for future work and focus in this study on processing conventional weighted MRI scans of T1, T2, and T2* contrasts to present a proof of concept of predicting microanatomical features from MRI.

3.2.2 Histology

Through collaborating with an external lab, we acquired three different histological stains from the same rat that we imaged for in-vivo and ex-vivo MRI acquisitions. Stains for microstructural features included Nissl histology, Myelin Basic Protein, and Iron labeling. These tissue labeling procedures were performed on slices cut from a 7.6 mm thick slab embedded in paraffin. Each 5 μm thick slice was stained according to staining scheme that was designed earlier to optimize the data acquisition and ensure uniformity in the order of slices across every stain. This scheme described a cycle of 26 slices which included 4 Nissl slices, 2 Myelin slices, 2 Iron slices, and 18 unstained slices. This cycle was repeated in the same order for the entire slab (15 times) and is summarized by Table 3. The unstained slices were also mounted and screened for multiple purposes. Firstly, observing all unstained slices before staining allowed us to evaluate the quality of the cutting and eliminate all damaged slices from the staining procedure, saving both time and cost. Secondly, having a continuous and uniform chain of slices (both stained and unstained) from the entire slab allowed us to employ our registration pipeline more efficiently. Since we have a good approximation of the slices positions across the z-axis based on our knowledge of the thickness of the slab and slices, we are able to narrow down the search window of the

registration algorithm and find the corresponding 2D MRI slice to the targeted histology slice in an optimized manner.

Table 3: Histology staining scheme

Slice number	Staining action
1	**Nissl**
2	unstained
3	**Myelin**
4	unstained
5	unstained
6	unstained
7	**Nissl**
8	unstained
9	unstained
10	unstained
11	**Iron**
12	unstained
13	unstained
14	**Nissl**
15	unstained
16	**Myelin**
17	unstained
18	unstained
19	unstained
20	**Nissl**
21	unstained
22	unstained
23	unstained
24	**Iron**
25	unstained
26	unstained

Even though iron histology was acquired, we discarded it from our analysis pipeline. This is due to the fact that the amount of iron detected by the staining in the rat brain we sliced was negligible. An example of a histology slice with scattered iron cells is shown in

Figure 5. The pink color is the background stain, while cells with iron should appear as blue in Prussian Blue stained slices (as shown in the right part of the figure).

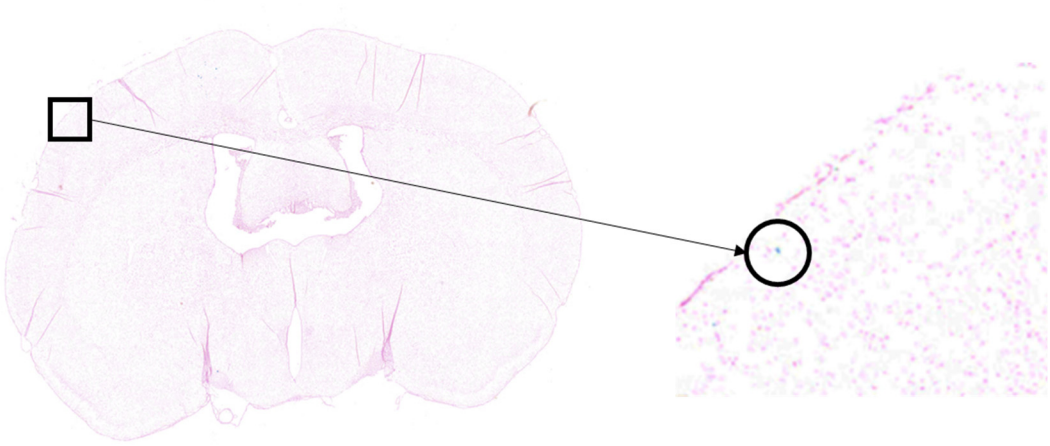

Figure 5: Example of slices stained for iron

3.2.3 Preprocessing

Before performing the registration and deep learning, we had to preprocess the data (MRI and histology) to ensure optimal results. First, we performed skull stripping for the in-vivo MRI data using a method described by [134]. To segment the brain, we initially register a reference volume with its brain mask to our in-vivo scan of interest. In this case, we used the ex-vivo scan of the same brain since it is the same brain and is already physically stripped from the skull and scalp. Using an affine transformation model, the mask of the reference scan is overlayed onto the in-vivo brain. This mask is then eroded and serves for the initialization of the brain extracting technique using level-sets segmentation. Through expansion, advection, and curvature terms, the level-set evolves towards the brain-skull border until all edges are matched [134]. However, we found that using the full volume (with skull, scalp, and surrounding structures) resulted in better registration. Therefore, we

implemented the registration pipeline using the full in-vivo scan, then used the calculated transformation on the other skull-stripped versions to use in the deep learning.

As for the histology, we down sampled the data to match it - approximately- with the in-plane resolution of MRI. This significantly reduced the computational time required to run an entire histology through the pipeline while maintaining satisfactory registration results. Following that, we processed the histology images to eliminate the noise and heterogeneity of the background. To automate the segmentation process, we assumed that the saturation values (from the HSV color model) of the background were low and that the brain formed one connected component in each histology slice. Therefore, we first removed the sporadic noise using Wiener filtering [135] then removed areas of low saturation values using thresholding. Finally, we only kept the largest single connected component in the image (the brain) using a popular morphology analysis technique of erosion, dilation and region filling [136].

To further improve the appearance of the slices before registering them to MRI, we removed the variability in intensity – a common artifact in histology images. This artifact shows as an inconsistency in staining across the histology slices and can be eliminated by histogram matching, which is mapping the intensity distribution of the slices to a reference distribution [136]. However, since not all anatomical structures appear on all slices, luminance distributions are different across slices and a global intensity histogram cannot be used. To account for this biological variability, we matched the histogram of slices with their neighbours using a sliding window rather than matching the entire dataset at once.

Finally, after registration, and even though the MRI and histology are of the same brain of the same animal, there will still be some regions that show on the MRI but not on the histology due to losses from cutting or tearing. These areas - especially if they are large relative to the training patch size - might mislead the learning algorithm and cause severe errors. Therefore, we eliminated those areas in the MRI that cannot be mapped to brain tissue in histology from the analysis.

3.3 Registration

Multimodal registration between histology and MRI is crucial for the feature learning process. If the anatomical features of the two modalities do not align well, the neural network could be misguided, and the results would be unreliable. To accurately align the different stains of histology to both in-vivo and ex-vivo MRI, we used the methodology and algorithm developed by Pilgram and Shmuel [137] and tailored it to our application. This pipeline describes a registration technique that finds and extracts 2D MRI slices that best correspond to 2D histology slices from a 3D MRI reference template and registers them well. As described in the literature review, using an intermediate image to register images that greatly vary in contrast and resolution is recommended. Therefore, we used ex-vivo images as the intermediate bridge for registering histology to in-vivo images. We first registered histology to ex-vivo MRI, then used these registered histology images and aligned them to in-vivo images using the pipeline described.

The first step in the pipeline developed by Pilgram and Shmuel is to localize the position of each 2D histology slice along the z-axis of the 3D MRI volume [137]. This was done using 3D affine registration in a coarse-to-fine strategy while taking the oblique angles

into account. To guide the process, registration quality measures were estimated at each scale level after predicting the reference MRI slice corresponding to a specific histology slice using diffeomorphic highly regularized demons' registration [138]. The metric used to assess the quality of the registration and guide its optimization included a combination of normalized Mutual Information and Mattes Mutual Information [139]. This metric is considered the gold standard for registration between multimodal images where intensity distributions are significantly different but general structures are similar. To optimize the parameters of the affine transformation, stochastic gradient descent optimization was implemented by the author [137]. Also, only random patches of pixels were used in the registration instead of the entire image to reduce the computational cost.

Further using the pipeline developed by Pilgram and Shmuel [137], several similarity measures were used to evaluate the quality of registration between the 2D multimodal images (i.e., histology and reference MRI slices) along the registration pipeline. After each stage, the weighted combination of two similarity metrics - the structural similarity index measure (SSIM) and the entropy correlation coefficient (ECC) - was calculated, with a maximum value of one for perfectly aligned images. SSIM [140] is calculated based on the difference in luminance, structure, and contrast between two images, while ECC is similar to normalized mutual information and is independent of individual values of entropy [141], [142]. The steps of the registration pipeline implemented on our dataset is summarized by the schematic illustration in Figure 6.

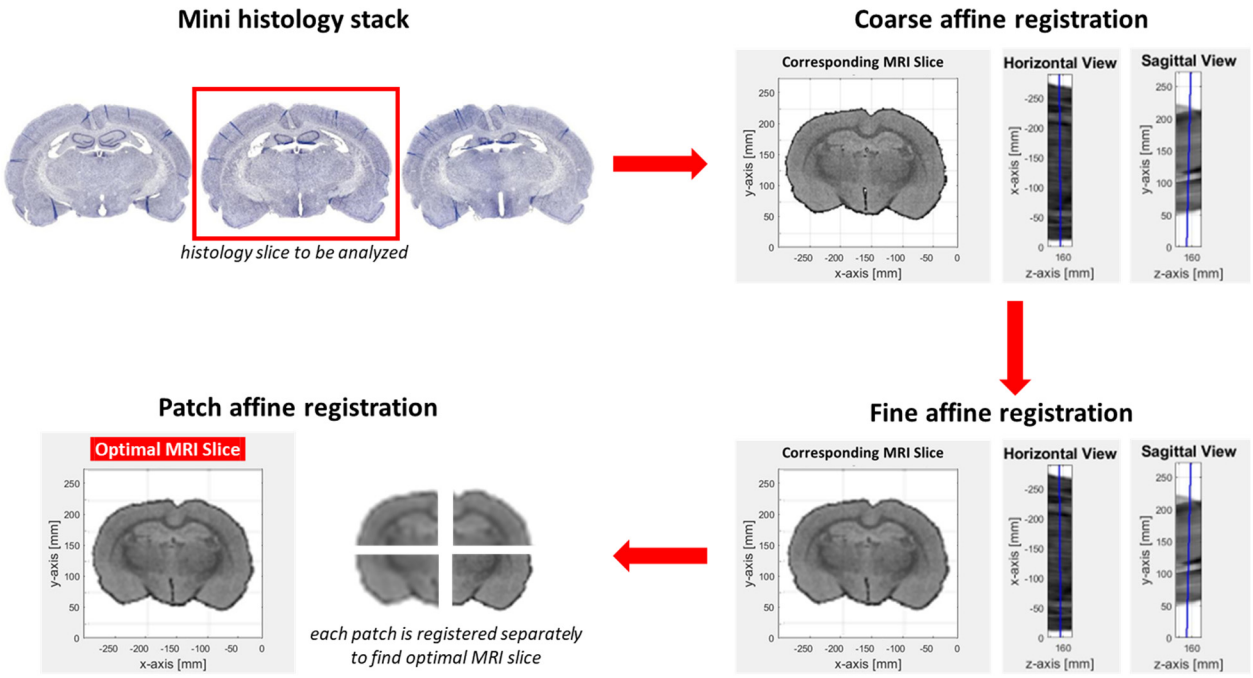

Figure 6: Schematic demonstration of the registration pipeline adopted from Pilgram and Shmuel's implementation [137]

First, the user saves all the histology slices in order (for example, anterior to posterior) and defines approximate values for the positions in millimeters of the first slice and the last slice along the z-axis. These defined values will be the basis for a search window with ±2 mm tolerance (or any other tolerance defined by the user). After selecting the 3D MRI volume from which we wish to extract 2D slices that correspond to the desired histology, the following steps are implemented to each histology image slice to extract its optimal reference MRI slice, as developed by [137]:

1- Build a 3D block of histology: three adjacent slices are selected to build an isotropic mini-stack, with the target histology slice to be aligned to the 3D MRI volume in the middle. These three slices are all registered to one another (with the middle slice being the reference) through affine registration. Following that, a second

registration operation is implemented using the highly regularized demons' transformation, thereby creating a smooth and compact 3D stack from these three adjacent images and removing any coarse deformations across them. Finally, linear interpolation is used to resample and generate isotropic images to match the resolutions of the 3D histology stack and the 3D MRI, with a minimum of five isotropic slices [137]. This now creates our 3D block of histology.

2- Extract a corresponding 3D block of MRI: based on the width of the search window and the pre-defined z-axis position of the target histology slice, a 3D block of MRI is extracted.

3- Register the 3D histology block to the 3D MRI block: using Elastix [143], the blocks from the two modalities are registered in three stages, from coarse-to-fine-to-patches:

- Coarse affine registration: the 3D histology block is first registered to the 3D MRI block and a spatial reference is estimated with approximated z-axis position and angles of rotation.
- Fine affine registration: the 3D histology block is then registered to the new 3D MRI block that was found in the previous step. However, the search window is halved to force a more accurate estimation of the z-axis position of the corresponding MRI slice and its rotation angles.
- Patches affine registration: Finally, the 3D histology block is divided into patches/sub-blocks of $<a \times b>$ size with up to 20% overlap allowed. Every histology sub-block is separately registered to its corresponding MRI sub-block and its registration quality is calculated based

on the combined metric of SSIM and ECC that was described earlier. We used the transformations from the sub-blocks that had the highest registration quality to obtain an optimal estimation of the z-axis position and rotation angles of the entire reference MRI slice for this specific histology slice.

Finally, since we are performing patch-based learning, we implemented a second round of non-linear registration between the patches of MRI and histology that will be fed to the deep learning pipeline, and employed the same metrics to assess their registration quality and eliminate all patches whose registration was unsatisfactory from the analysis (See Chapter 4 for the evaluation of unsatisfactory registration).

3.4 Deep learning framework

One type of neural network architectures that grew popular recently in image-to-image translation applications are conditional general adversarial networks (cGANs). A typical non-conditional GAN consists of two sub-networks competing in a zero-sum game: a generator ***G***, and a discriminator ***D***. The goal of ***G*** is to learn a mapping between a random noise ***n*** and a target image ***t*** that will allow it to generate a plausible fake replica of the target image. ***D*** on the other hand is trained to distinguish a fake images ***G(t)*** from the real image ***t***. The learning process of both neural networks is simultaneous, where ***G*** is trying to learn the best mapping that will trick ***D***, and ***D*** is trying to detect the fakes that are generated by ***G***. The following equation explains how the adversarial loss is used by **G** and **D**:

$$L_{GAN}(G, D) = E_t[logD(t)] + E_n[\log\Big(1 - D\big(G(n)\big)\Big)]$$

The goal is for ***G*** to minimize the adversarial loss and for ***D*** to maximize it. High-frequency information in the image can be modeled by the generator by the help of

adversarial loss [132]. The two networks will keep competing in this game until convergence occur when ***G*** has learned a reliable mapping to the point that ***D*** is no longer able to detect the fake images generated by ***G***.

Our aim is to learn and predict histology from MRI images. Since both images represent the same general anatomy and structures of the same subject, the goal is to map from a source image ***s*** (rather than noise) to a target image ***t***, which has a statistical dependency and shares a visual common ground with ***s***. Therefore, it is necessary to condition the prediction of the GAN by incorporating the knowledge of s [144]. This is called cGANs and can be represented by the following adversarial loss:

$$L_{cGAN}(G,D) = E_{s,t}[logD(s,t)] + E_{s,n}[\log\Big(1 - D\big(s, G(s,n)\big)\Big)]$$

One common architecture for cGANs that was proposed in 2018 by Isola et al. [145] and has been used in numerous image synthesis applications ever since is the pix2pix architecture. Based on the assumption that the MRI and histology are very well registered to one another, we adopted the pix2pix framework. In previous implementations, mixing a traditional loss – such as L2 or L1 distance – with the GAN objective yields better results [146]. Therefore, in addition to fooling the discriminator, the generator's new task will be to minimize the distance between the prediction and the ground truth. We used L1 distance since it does not enforce blurring to the level L2 does:

$$L_{L1}(G) = E_{s,t,n}[\|\mathrm{t} - \mathrm{G(s,n)}\|_{L1}]$$

Therefore, the final objective of our architecture is:

$$L_{MRI2hist} = \arg\underbrace{min}_{G}\ \underbrace{max}_{D} L_{cGAN}(G,D) + \lambda L_{L1}(G)$$

Where λ is the factor determining the effect of the L1 loss compared to the adversarial loss.

Generator architecture: In our implementation, we adopted the infamous UNet architecture – a variant of an encode-decoder network – as the generator. A typical encoder-decoder network consists of an encoder module that down-samples the image input in a series of convolutional layers to extract features, followed by a decoder module that reverses the operation and up-sample the features back to the starting spatial dimensions. Since we are translating MRI to histology and vice versa, we design our generator according to the assumption that the images represent similar structures and are well registered to one another. The typical encoder-decoder structure forces all the information to flow in one direction across all layers, including the reversing point between the encoder and decoder. However, in our case, there are a lot of low-level features that are necessary to be shared between the encoder and the decoder, such as the existence of specific substantial structures (corpus collosum, ventricles, etc.) or the location of substantial edges. The solution to this, as suggested by Ronneberger et al. [147] in their implementation of U-Net, is to use skip connections between the encoder and the decoder. This simply means that connections are made between each layer i in the encoder and layer $n-i$ in the decoder (n being the number of total layers), and the channels from those two layers are concatenated together as shown in Figure 7. For both the encoder and the decoder, we used series of convolutions, batch normalization, and ReLU activation layers. Following the original implementation of pix2pix, we used a 4x4-pixel sized filters for all convolutions and a stride of 2. In the encoder, the convolutions were down-sampling the images by halves at each step and all the ReLUs used

were leaky at a 0.2 slope. In the decoder, the convolutions up-sampled by at each step by a factor of 2 and all ReLUs used were not leaky.

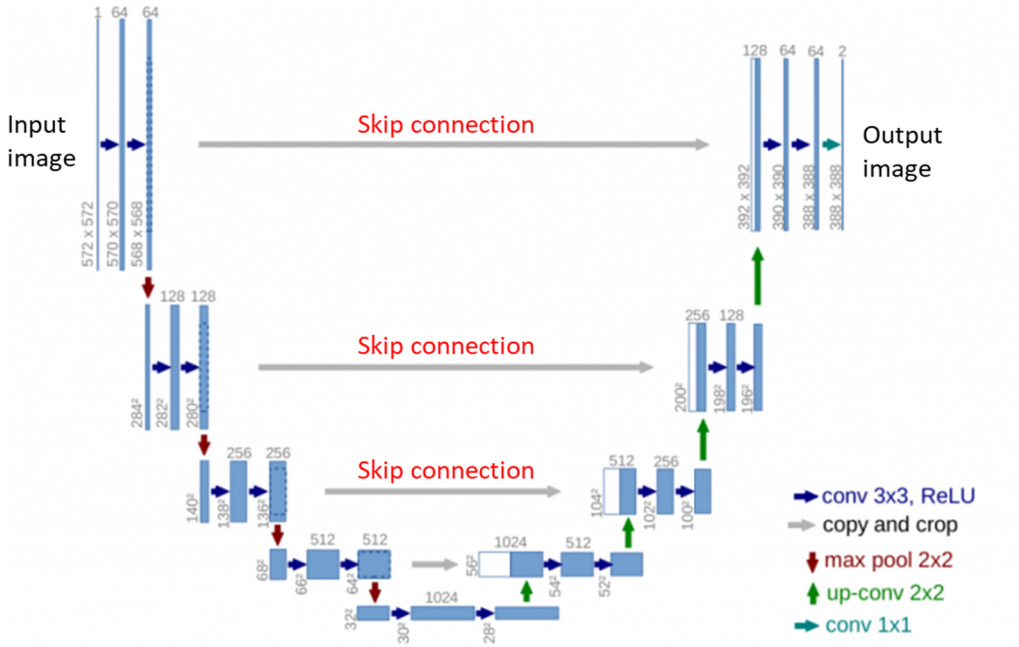

Figure 7: Demonstration of a standard U-Net architecture with skip connections [147]

Discriminator architecture: Conventional L1 and L2 loss encourage blurriness in image-based applications due to their favoring of mean values, failing to capture high-frequency features [148]. However, since L1 adequately captures low-frequency features and is less blurry than L2, we worked around its poor high-frequency capturing by using patches in the discriminator. In other words, we are forcing the cGAN discriminator to only train on modelling high-frequency features, while also maintaining the accuracy of low-frequency features through using L1. This means that the discriminator will only assess patches of the image and predict if it is real or fake. The results from all patches across the image are averaged and the final prediction of the discriminator is made. This type of discriminator is named a PatchGAN discriminator [149]–[151]. The structure of the

discriminator follows a repeated “convolution layer –> batch normalization layer –> leaky ReLU layer” sequence.

Optimization: We used Adam optimization and a learning rate of 0.005. Based on the implementation by [152], we alternated between the discriminator and the generator in the gradient descent steps. Also, as suggested by the pix2pix original paper [145], dividing the objective of the discriminator by 2 in the optimization will slow the learning rate of the discriminator in comparison to the generator and will improve the results. Figure 8 shows a simplified schematic of the pipeline.

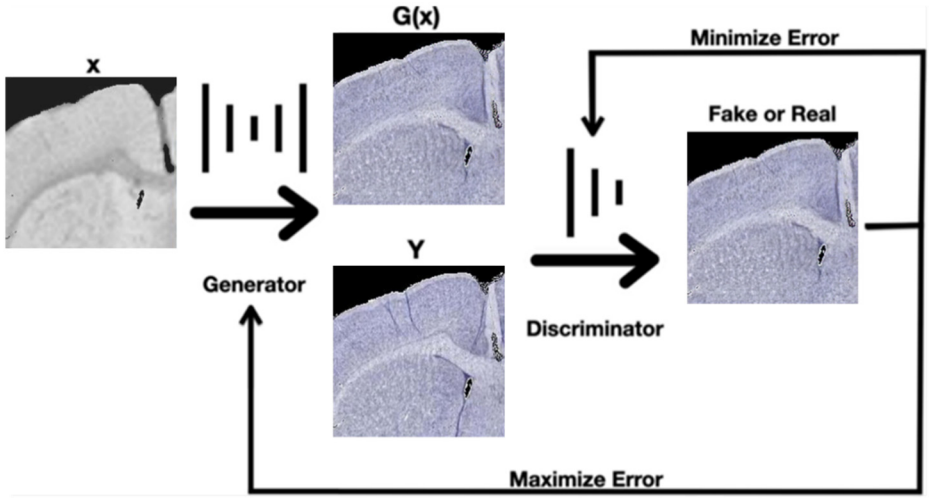

Figure 8: cGAN pipeline

3.5 Evaluation Metrics

Numerous metrics have been employed and discussed in the literature of previous image-to-image translation applications to assess the performance of the neural network on test datasets. Typically, a test dataset is created as a subset of the original dataset and is never seen by the neural network during the training process. This allows for an objective evaluation of the accuracy of the prediction and the bias/variance of the model. In our case,

we randomly split our data into 90% training data, and 10% testing data. We implemented a 5-fold cross-validation scheme on our training data (using 20% of the training data each fold). Since our prediction task is visual, we decided to employ both qualitative and quantitative evaluation methods. The following is a description of those methods:

Superimposed Edges: This is a qualitative method of visually assessing how the predicted image is similar to the ground truth image when overlaid over it. To implement this, edges are first detected on the predicted image using the Canny edge detection technique [153]. This technique uses a Gaussian filter derivative to calculate the gradient of the image, followed by a search for local maxima of the gradient to find the edges. Two thresholds are used to ensure that both strong and weak edges are detected, which makes the Canny technique less sensitive to noise than other methods, and more capable of catching weaker but real edges.

Peak signal-to-noise ratio (PSNR): This quantitative evaluation metric assesses the differences in intensities between a predicted image and its corresponding ground truth. Higher values of PSNR reflect a closer range of intensities between the two images. The following function is used to calculate PSNR:

$$PSNR = 10log_{10}\left(\frac{Peak\ Value}{MSE}\right)$$

The $Peak\ Value$ is the peak signal level in the image. For instance, if the range of intensities in the image is [0, 255], the $Peak\ Value$ will be defined as 255. The Mean Squared Error (MSE) is calculated pixel-by-pixel between the predicted image and the ground truth.

Structural Similarity Index (SSIM) [140]**:** This is one of the most common metrics to holistically evaluate the similarity between the predicted image and its ground truth through computing the perceptual distance between the two images. The index is a value

between -1 and 1, where higher SSIM values indicate greater similarities between the structure, contrast, and luminance of the two images, and vice versa. The final metric is a combination of those three elements as is shown in the following equation:

$$SSIM(x,y) = [structure(x,y)]^a . [contrast(x,y)]^b . [luminance(x,y)]^c$$

The notions "contrast", "structure", and "luminance" are based on findings and observations of the authors [140]. For example, they concluded that averaging over all the pixel values of the image is a good descriptive of the luminance. Therefore, if the two images differ greatly in brightness, the luminance term will drop. These features are mathematically represented as follows:

$$structure(x,y) = \frac{C_1 + \sigma_{xy}}{C_1 + \sigma_x \sigma_y}$$

$$contrast(x,y) = \frac{C_2 + 2\sigma_x \sigma_y}{C_2 + {\sigma_x}^2 + {\sigma_y}^2}$$

$$luminance(x,y) = \frac{C_3 + 2\mu_x \mu_y}{C_3 + {\mu_x}^2 + {\mu_y}^2}$$

Where σ_x and σ_y are the standard deviations, σ_{xy} is the cross-covariance, and μ_x and μ_y are the means for images x and y. C_1, C_2, C_3 are constants used to avoid instability when certain parameters are close to zero.

Chapter IV – Experiments and Results

4.1 Experiments Scheme

We are interested in analyzing the relation between the different MRI parameters (T1, T2, and T2*) with microanatomical features based on Nissl staining and relative Myelin content presented in histological data. According to the literature, several microstructural variations could affect MRI measures. Therefore, combining multiple MRI contrasts is likely necessary for extracting meaningful and quantitative microstructural or features from the brain tissue. This will also allow us to distinguish microanatomical features that contribute the most to the MR signal from those that have insignificant effects. Therefore, to study this multiparametric relationship, we conducted a number of different experiments to predict individual microanatomical features from individual MRI contrasts or a combination of contrasts. These experiments involve training separate deep learning models that we later quantitatively assess to determine the best predictor (or predictors) of a specific tissue feature. Table 4 shows our experimental scheme including unimodal, bi-modal, and multimodal experiments. Since we acquired T1 and T2, and T2* with several TRs and TEs, we selected the individual TRs and TEs with best contrast and SNR. Namely, all T1 images selected for this analysis were acquired with a TR of 1,200 ms, T2 images were acquired with a TE of 19.27 ms, and T2* images with a TE of 7.56 ms. Also, since we are using single TR and TE acquisitions for analysis, it is important to mention that these are now weighted acquisitions and not quantitative maps. Therefore, our use of the terms "T1, T2, and T2*" from this point onwards will be referring to these weighted acquisitions. Moreover, even though T2* images were included in the data we acquired in-vivo, we did not use them in the analysis since they were too noisy to be of any use in the predictions. This left us with 20

different experiments for our three different tissue labels. We trained all our models through remote computing on Compute Canada's cluster Cedar using both CPU and GPU cores. For each experiment, we used 3,000 images of 400x400 patches of different magnifications (all rescaled to the same size) of MRI slices and their corresponding histology slices. Meaning, the training dataset consisted of images representing the full brain, as well as images representing patches from the brain at different magnifications (~10%, 15%, and 20% of the brain size). The goal behind this is to ensure that the network is being trained on all low-level and high-level features at different magnifications. Even though the size of our training set is not exceedingly large and different parameters were optimized for faster processing, it still took ~12 hours to compute 500 epochs of each model. Therefore, using Compute Canada's cluster enabled us to perform parallel computations of different models simultaneously and accelerate the overall process. For each experiment and deep learning model, we qualitatively assessed the quality of predictions using superimposed edges on testing dataset. We also calculated and reported the quantitative evaluation metrics of PSNR, SSIM, and MSE.

Table 4: Summary of all validation experiments to predict histology from MRI

Histology label	MRI environment	Unimodal	Bi-modal	Multimodal
Nissl	Ex-vivo	T1	T1 & T2	T1 & T2 & T2*
		T2	T2 & T2*	
		T2*	T1 & T2*	
	In-vivo	T1	-	T1 & T2
		T2		
Myelin	Ex-vivo	T1	T1 & T2	T1 & T2 & T2*
		T2	T2 & T2*	
		T2*	T1 & T2*	
	In-vivo	T1	-	T1 & T2
		T2		

4.2 Registration results and evaluation

The registration pipeline described earlier was implemented to achieve adequate alignment between MRI and histology images, which is necessary for facilitating the learning process and ensuring accurate predictions. We also used the evaluation metrics discussed in the previous chapter to assess and improve the registration quality. Figure 9 presents an example of the registration result for a histology slice with its corresponding 2D MRI slice that was identified and extracted from the 3D volume using the methodology described in the previous chapter. Figure 10 demonstrates a qualitative assessment of the registration of the same slices using superimposed edges. The red edges were computed from the histology image. They can be seen tracing the external edges of the brain in the histology image. When superimposed on the MRI image, they almost trace the same shape. This method can also be used to visually track the alignment quality of certain regions or structures, such as the white matter, ventricles, or corpus collosum.

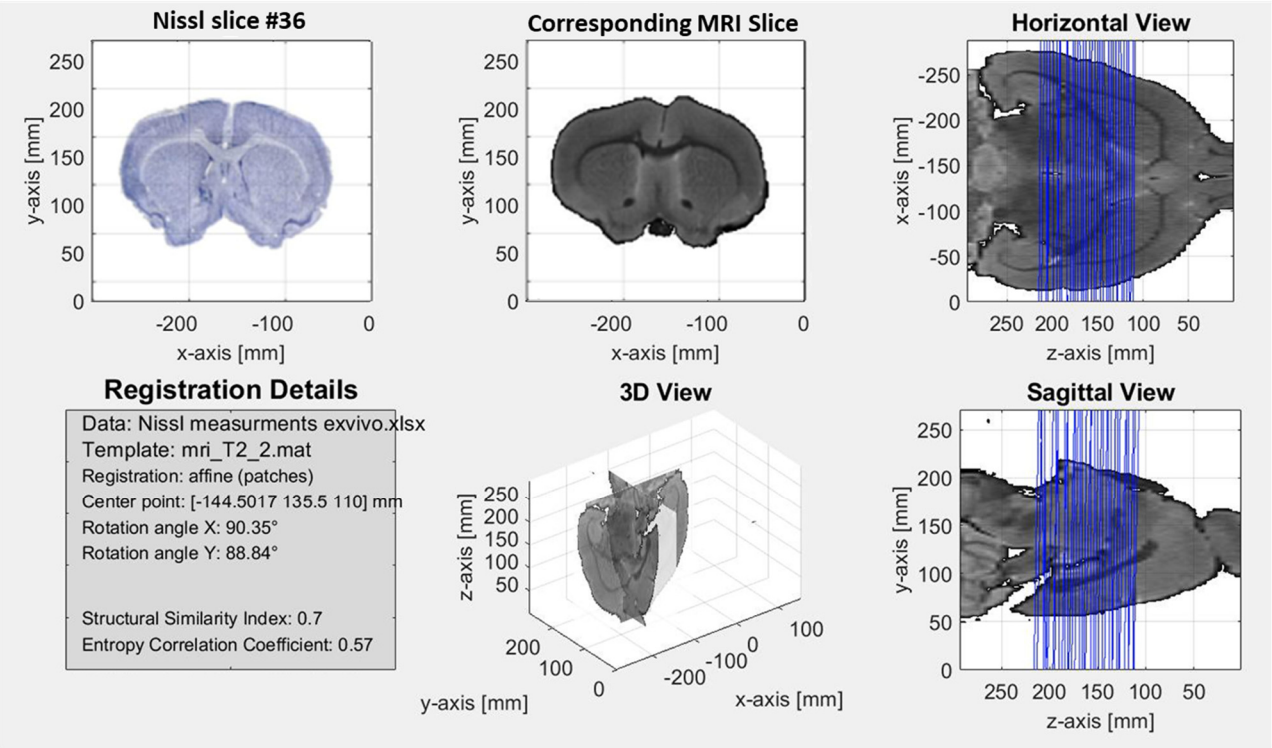

Figure 9: Registration example between a Nissl-stained histology image and the corresponding T2 MRI image localized and extracted by the software developed in [137]

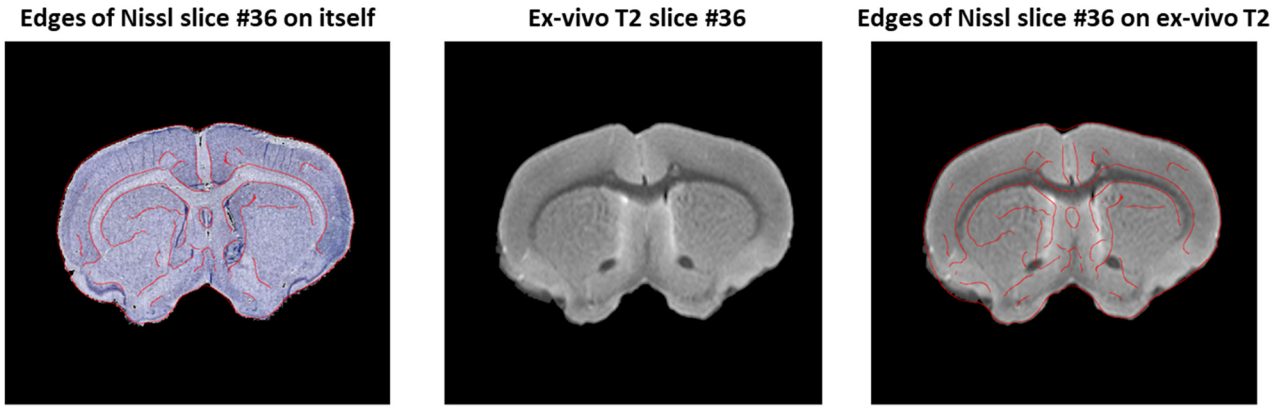

Figure 10: Qualitative assessment of the registration of a Nissl slice and its corresponding ex-vivo T2 MRI slice

Based on assessments of the registration quality and on the fact that some histology slices were significantly severed during the cutting or staining process, we had to eliminate

a few slices (~5% of the data) as shown in Figure 11. All slices that were included in the analysis had the pattern expected from previous MRI and histology. In other words, distortions or tears were minimal.

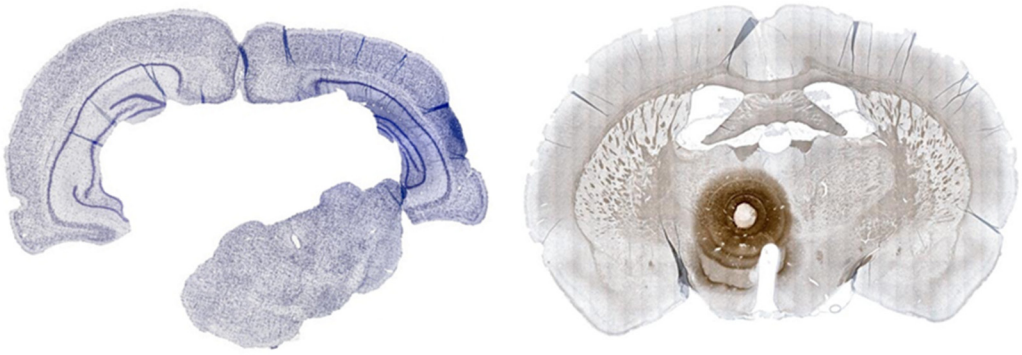

Figure 11: Examples of severely damaged histology slices that were eliminated from our analysis. The slice on the left is a Nissl slice that got torn in the cutting process, while the slice on the right is a Myelin slice that had a staining artifact.

In addition, since we perform patch-based learning, we implemented a second round of non-linear registration between patches and employed the same metrics to assess their registration quality and eliminate all patches whose registration is unsatisfactory, before feeding them to the deep learning pipeline. This second round of patch registration is different than the patch registration described in the methodology section that was used to determine the most accurate localization of the corresponding 2D MRI slice. Figure 12 demonstrates examples of satisfactory and unsatisfactory patch registration. The number of patches eliminated in this process varied for each experiment, but was always <5% of the dataset size.

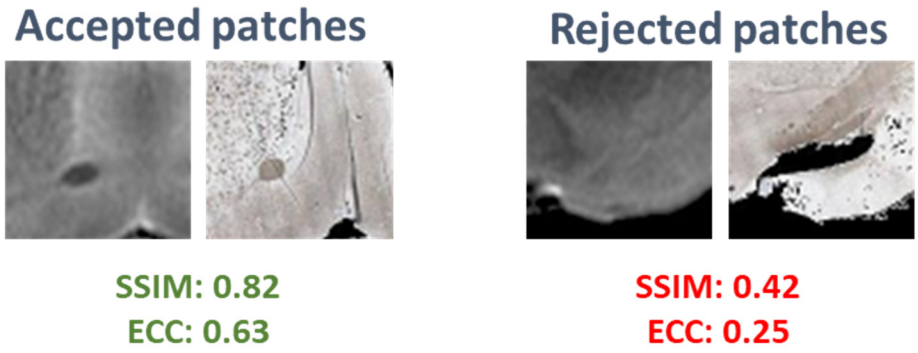

Figure 12: Example of the second round of registration of patches to eliminate patches that were poorly registered.

Below is a summary of the quantitative evaluation metrics for the registration quality of all the pairs that were included in our analysis. We used the same metrics that were explained in the methodology section. The average and standard deviation across each set of images is reported in Table 5. The best and worst experiments according to each metric are highlighted in green and red respectively.

Table 5: Summary of registration quality evaluation for all different registration pairs

Histology label	MRI environment	MRI contrast	SSIM		ECC	
			Avg	*SD*	*Avg*	*SD*
Nissl	Ex-vivo	T1	0.71	0.091	0.56	0.033
		T2	0.68	0.079	0.59	0.031
		T2*	0.74	0.075	0.60	0.036
	In-vivo	T1	0.67	0.132	0.46	0.088
		T2	0.58	0.112	0.41	0.098
Myelin	Ex-vivo	T1	0.77	0.042	0.66	0.028
		T2	0.75	0.058	0.62	0.030
		T2*	0.80	0.044	0.66	0.024
	In-vivo	T1	0.64	0.098	0.49	0.075
		T2	0.61	0.093	0.48	0.081

Validation experiment – Predicting different anatomical views:

Before running our pipeline on the real dataset that we are interested in analysing, we need to validate the concept and assess the learning capabilities of our algorithm on dummy data. Since we are significantly constrained by the availability of data, working with images from a single animal, we needed to verify and prove that our algorithm was not overfitting and was actually learning the local features between MRI and histology. To do that, we aimed to train and test the models on two different anatomical views. For instance, we train the model using coronal slices of MRI and histology, then test using horizontal slices and observe the results. If the model succeeds to map the features from a horizontal MRI slice to a horizontal histology slice, then this indicates that the underlying learning mechanism is functional and meaningful. However, since it is not possible to obtain horizontal histological slices from the same animal after cutting the brain along the coronal axis, we could not perform this validation due to the lack of ground truth data to evaluate the predictions. Since the purpose behind this is merely validating the concept, we tested the algorithm through an alternative experiment of predicting ex-vivo images from in-vivo images. Since the MRI data is acquired as 3D volumes, it is possible to extract both coronal and horizontal slices from the same acquisition. As mentioned, the goal is to be able to predict horizontal ex-vivo images using a model that was trained using only coronal slices and features from the coronal view. We used the same registration pipeline and same deep learning hyperparameters to conduct this experiment. Figure 13 demonstrates the registration result for one of the slices.

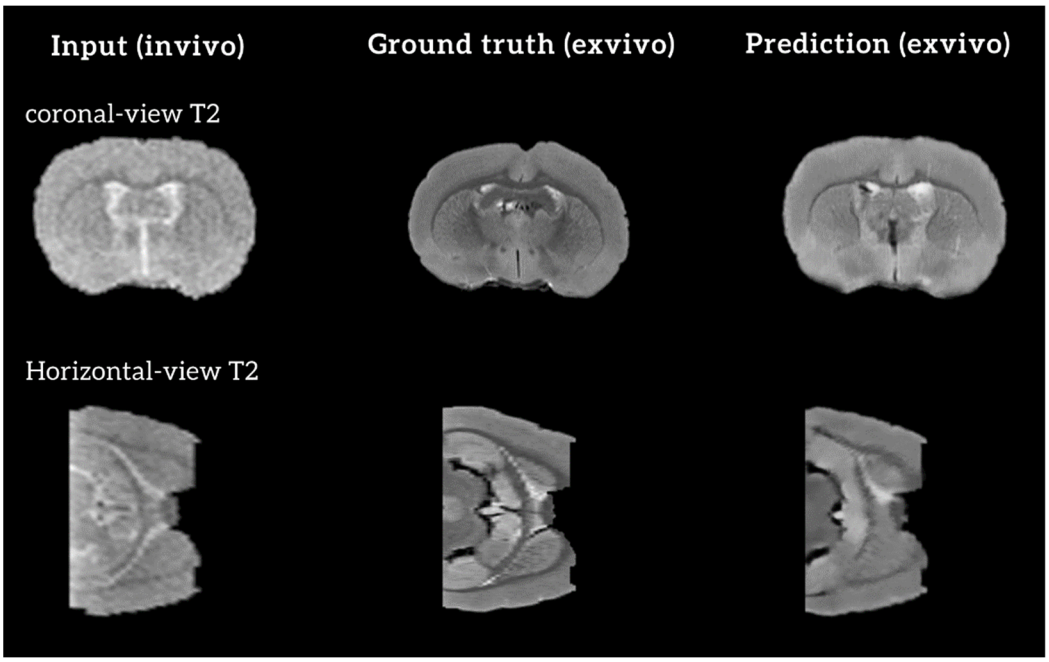

Figure 13: Predicting ex-vivo horizontal slices using a model trained only on coronal slices

We also employed the same evaluation metrics to assess the quality of predictions. Table 6 presents a summary of the findings.

Table 6: Quantitative evaluation of validation experiment (in-vivo to ex-vivo)

Slice orientation	**PSNR**		**SSIM**	
	Avg	*SD*	*Avg*	*SD*
Coronal	23.201	1.113	0.64	0.0612
Horizontal	15.673	3.894	0.56	0.298

4.3 Uni-modal experiments

In this set of experiments, we are interested in investigating which individual MRI contrast best predicts each microanatomical feature. To the best of our capabilities, we ensured the preprocessing of the datasets, the training environment, and the network hyperparameters were identical and consistent throughout all experiments. Table 7 lists all 10 experiments along the average quantitative evaluation on the test set consisting of 50 images of varying magnifications, resolutions, and regions of the brain. The best and worst experiments according to each metric are highlighted in green and red respectively.

Table 7: Summary of quantitative evaluation of predictions for uni-modal experiments

Histology Label	MRI environment	MRI contrast	PSNR		SSIM	
			Avg	SD	Avg	SD
Nissl	Ex-vivo	T1	21.873	2.748	0.653	0.121
		T2	19.891	3.776	0.549	0.202
		T2*	22.770	2.494	0.659	0.165
	In-vivo	T1	18.889	3.012	0.527	0.207
		T2	18.296	3.009	0.516	0.199
Myelin	Ex-vivo	T1	20.393	2.175	0.655	0.118
		T2	21.000	2.412	0.658	0.160
		T2*	21.999	2.475	0.690	0.124
	In-vivo	T1	19.028	3.491	0.569	0.220
		T2	18.372	3.573	0.500	0.193

Figure 14 and Figure 15 demonstrate different examples of predictions for Nissl and Myelin labels from different MRI contrasts. It is important to mention that the results from in-vivo MRI slices appear slightly different from ex-vivo MRI slices because the imaging environment and parameters were different. Importantly, also the tissue has different microstructural features. This however had no impact on the training process since both scans were used separately.

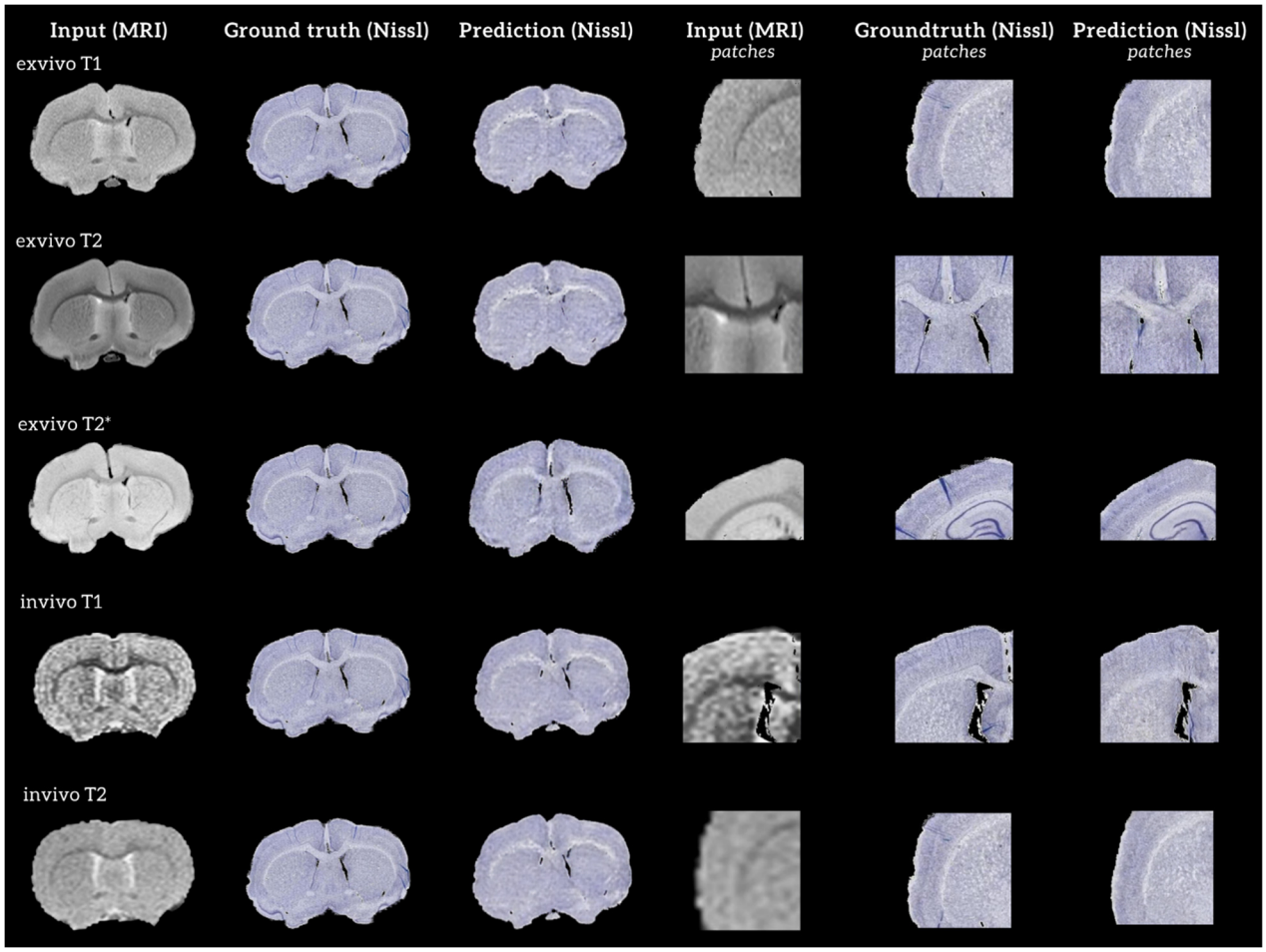

Figure 14: Unimodal evaluation to predict Nissl from different individual MRI contrasts

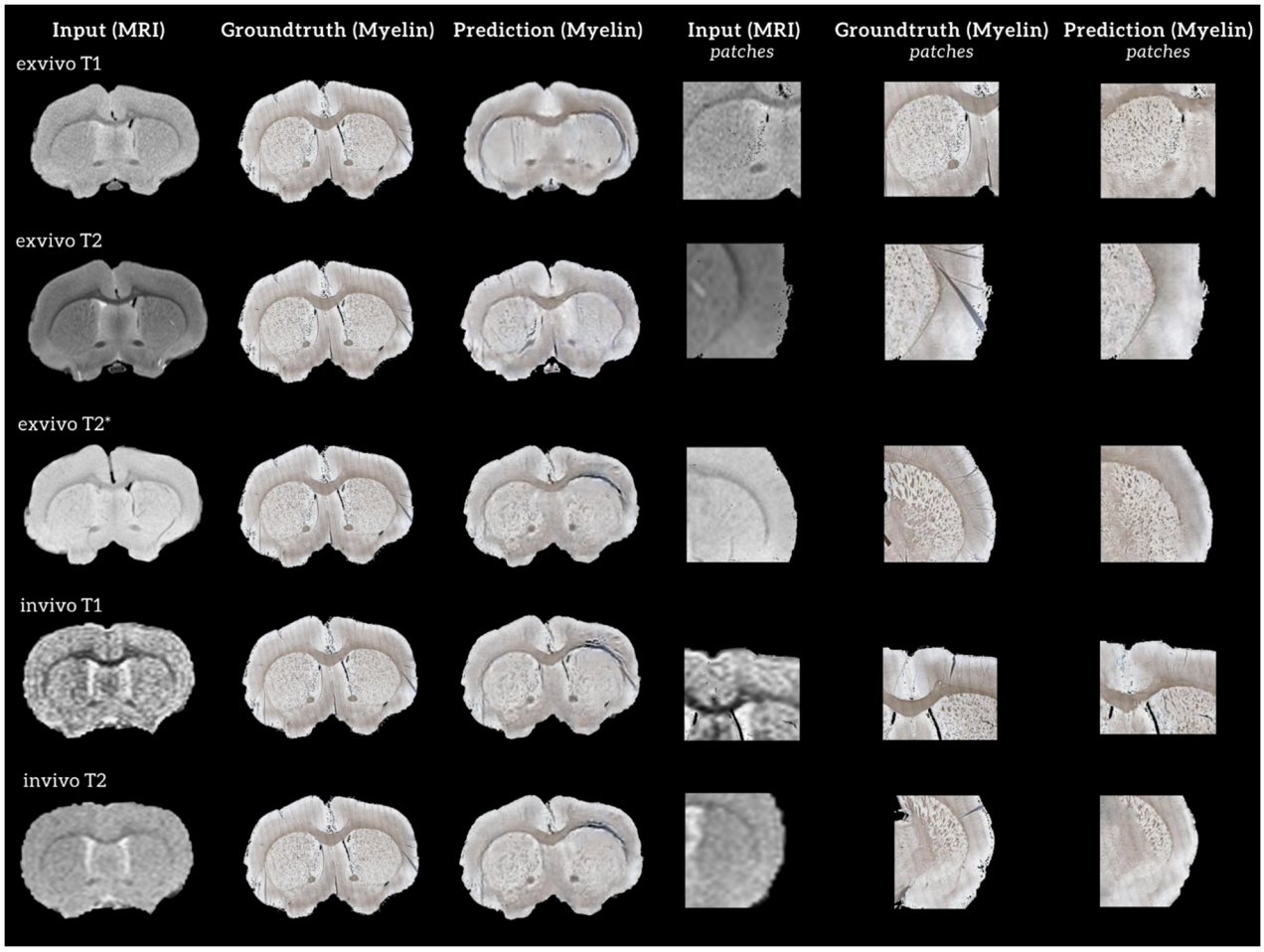

Figure 15: Unimodal evaluation to predict Myelin from different individual MRI contrasts

4.4 Bi-modal experiments

In this set of experiments, we wanted to evaluate the impact of combined MRI contrasts in predicting individual histology labeled images. Since we have three MRI measurements (T1, T2, and T2*) we implemented a leave-one-out scheme to find out which MRI contrast had the least impact in predicting each histological feature. Since we eliminated in-vivo T2* scans from all experiments and we only have two in-vivo scans, we will be reporting the in-vivo experiments with multi-modal experiments. This leaves us with a total of 6 experiments for Nissl and Myelin. Table 8 names the different experiments and presents

a summary of the average quantitative evaluation using PSNR and SSIM on the same test set used for unimodal experiments. The best and worst experiments according to each metric are highlighted in green and red respectively.

Table 8: Summary of quantitative evaluation of predictions for bi-modal experiments

Histology label	MRI environment	MRI contrasts	PSNR		SSIM	
			Avg	SD	Avg	SD
Nissl	Ex-vivo	T1 & T2	23.021	1.976	0.726	0.110
		T1 & T2*	23.997	2.843	0.777	0.108
		T2 & T2*	24.344	1.332	0.819	0.083
Myelin	Ex-vivo	T1 & T2	24.001	1.102	0.816	0.077
		T1 & T2*	23.872	1.366	0.834	0.099
		T2 & T2*	23.749	2.018	0.714	0.137

Figure 16 and Figure 17 demonstrate examples of predictions for full brain and magnified patches of Nissl and Myelin. For demonstration purposes, we split the MRI full brain slice in two halves to show the two contrasts involved in each bi-modal experiment.

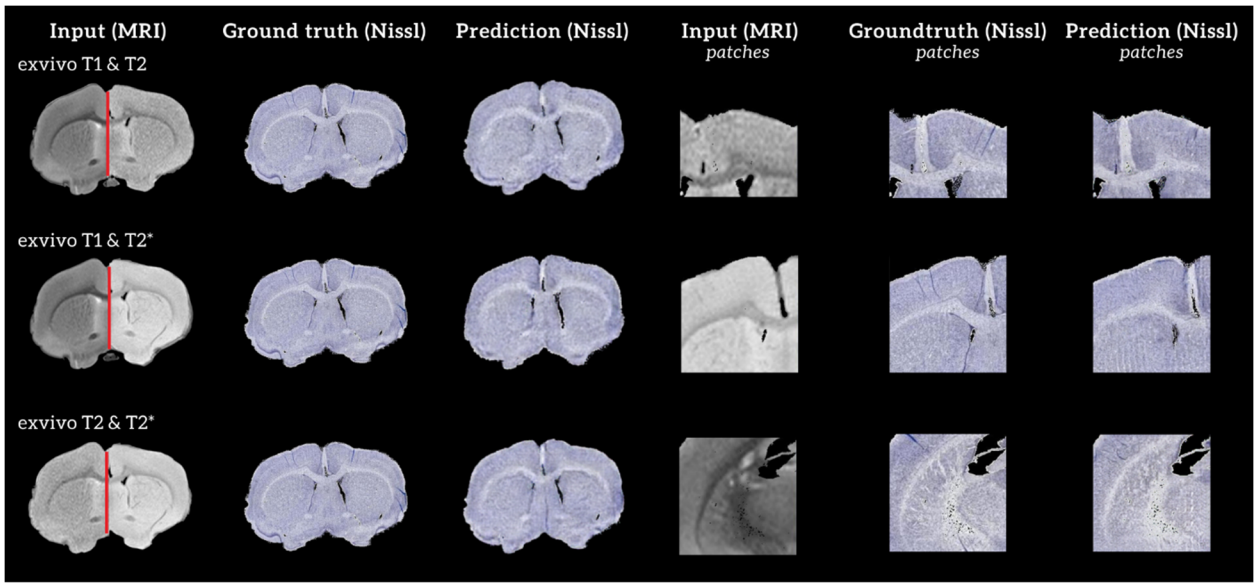

Figure 16: Bi-modal evaluation to predict Nissl from different combinations of MRI contrasts. For demonstration purposes, we split the MRI full brain slice in two halves to show the two contrasts involved in each bi-modal experiment

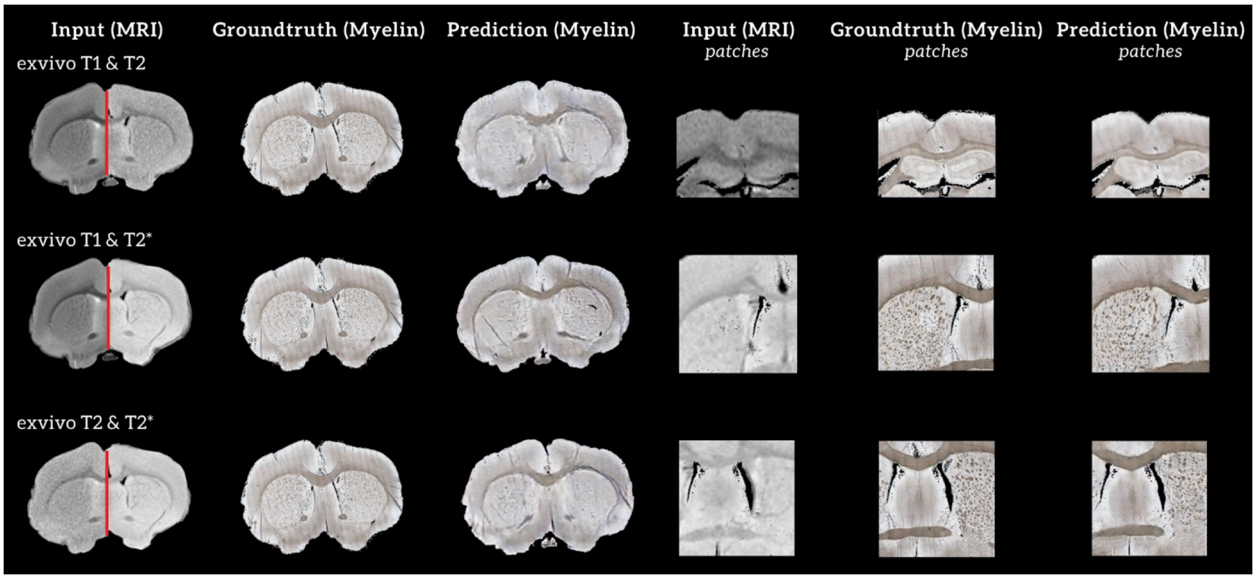

Figure 17: Bi-modal evaluation to predict Myelin from different combinations of individual MRI contrasts

4.5 Multi-modal experiments

Similar to our investigation with bi-modal MRI experiments, these set of experiments take a look at the effectiveness of using all MRI contrasts (T1 and T2 for in-vivo and T1, T2, and T2* for ex-vivo) to predict specific microanatomical features. Table 9 lists all four experiments and reports their average PSNR and average SSIM measures along with the standard deviation from each experiment. The best and worst experiments according to each metric are highlighted in green and red respectively.

Table 9: Summary of predictions quantitative evaluation for multi-modal experiments

Histology label	MRI environment	MRI contrasts	PSNR		SSIM	
			Avg	*SD*	*Avg*	*SD*
Nissl	Ex-vivo	T1 & T2 & T2*	25.375	1.929	0.813	0.083
	In-vivo	T1 & T2	21.009	2.976	0.557	0.130
Myelin	Ex-vivo	T1 & T2 & T2*	27.161	1.893	0.822	0.080
	In-vivo	T1 & T2	20.811	1.993	0.594	0.089

Figure 18 and Figure 19 demonstrate examples of predictions for full brain and magnified patches of Nissl and Myelin. As we did with bi-modal experiments - for visual demonstration - we used red lines to divide the full brain slice to segments of the MRI contrasts involved in each experiment.

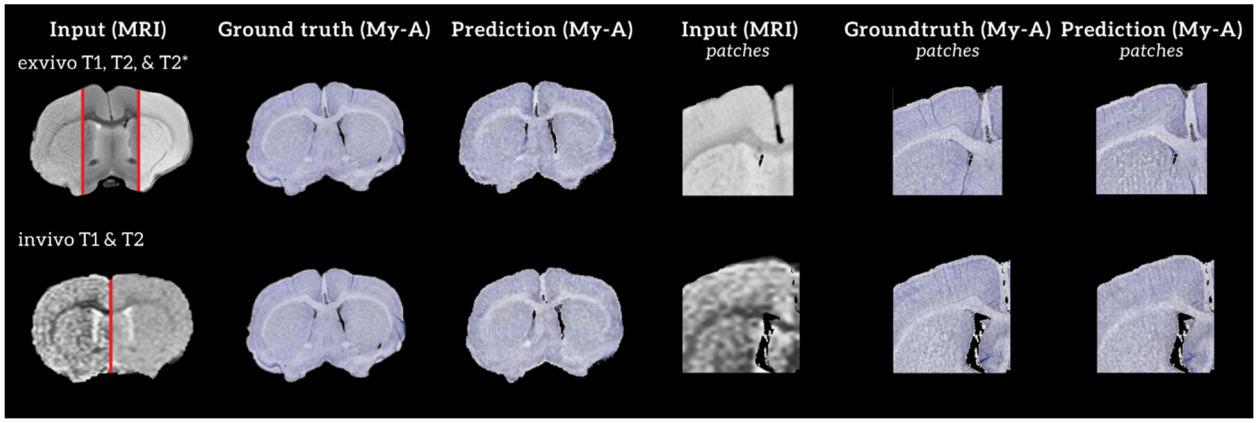

Figure 18: Multi-modal evaluation of predicting Nissl from all combinations of individual MRI contrasts

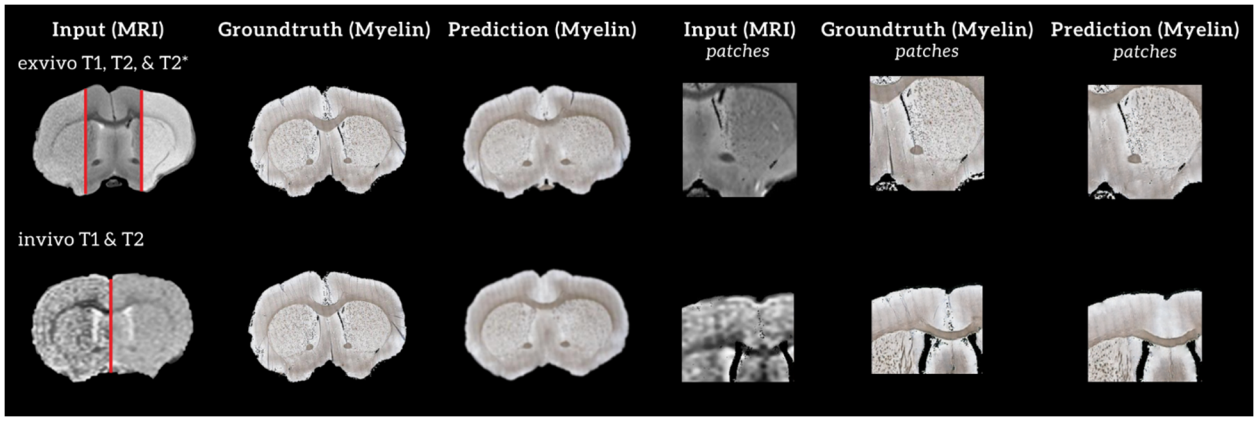

Figure 19: Multi-modal evaluation to predict Myelin from all combinations of individual MRI contrasts

Chapter V - Discussion

Numerous clinical applications and research opportunities in neuroscience could be made possible from the ability to estimate and predict microanatomical features noninvasively from in-vivo scans. This will enable more accurate comparisons between different individuals that track subtle microstructural differences. For instance, several studies have highlighted the correlation between age and changes in quantitative parameters such as MT, T1, and T2 [6], [84], [154]. These changes in quantitative MR parameters were also correlated with a corresponding increase in iron concentrations and demyelination effects related to age as shown in Figure 2. Moreover, in-vivo accurate estimation of microanatomical features can also help detect and monitor traumatic injuries and their consequences as well as degenerative processes. A spinal cord injury was observed by Freund et al. [155] and they recorded rapid morphological changes in the brain within 3 months of the injury. These changes were also expressed by the MT and T1 signals and suggested that myelinated axons were in retrograde degeneration. Another study by Deppe et al. [100] used DTI imaging to detect changes in white matter in Alzheimer patients. Iannucci et al. [156] discussed how MS patients could be potentially better monitored using a multiparametric MR approach. It is therefore clear that the ability to directly link quantitative MR parameters with microstructural features will have a wide impact on neurological studies and assessments.

Accurately estimating microstructural features in a standardized manner and relating them to MRI parameters opens another possibility for systematic comparisons between structure and the underlying functional activity. A recent study performed by Helbling et al.

[157] demonstrated how magnetoencephalography (MEG) measured functional response amplitudes can be predicted from qMRI maps that are sensitive to myelin. This shows how estimating features of the tissue in-vivo could be an important step in elucidating the structural-functional relationship in the brain at a microanatomical level.

The task of studying the relationship between microanatomical features and MRI parameters is primarily concerned with in-vivo MRI measurements. This is because the post-mortem brain is known to experience certain chemical and biological alterations that influence the quantifiable measurements of microanatomical features we are interested in. Therefore, the post-mortem ex-vivo brain is not expected to accurately reflect the biomarkers of interest for different pathologies and is subject to unquantifiable variations. However, there are several imaging constraints that limit the ability of in-vivo MRI acquisitions to correlate with microanatomical structures in an accurate and quantifiable manner. To start with, we have established earlier that histological imaging provides a much higher resolution (on the order of 1μm) compared to MRI (0.1-5 mm) and is therefore the current gold-standard for analyzing microanatomical features and structures. In addition, in-vivo and ex-vivo MRI acquisitions also provide a different range of resolutions. For instance, our in-plane resolution for in-vivo T2 was 200x200 μm compared to a 100x100 μm for ex-vivo T2, and the slice thickness for in-vivo T2 was 400μm compared to 300μm for ex-vivo T2. This is controlled by the fact that a living animal cannot be put under anesthesia or held steadily for too long, therefore limiting the scan time. Certain biological processes like heart beat and respiration also affect the quality of the in-vivo scans. Therefore, even though in-vivo scans are theoretically a better predictor of quantitative microanatomical measurements, its relatively poor resolution limits its utility in deep learning applications of

this kind. Thus, as a proof of concept, we included both in-vivo and ex-vivo scans in our analysis to explore and present the potentials of our methods. It is also important to mention that it is possible to acquire in-vivo scans with better resolution, either through optimizing the scanning parameters or through using a different scanner. This is an area of future work and improvement since better in-vivo acquisitions is expected to lead to more reliable conclusions about the relationship between MRI parameters and histological features.

5.1 Challenges and limitations

5.1.1 Data availability

There were several challenges and limitations faced while designing and implementing our deep learning framework to learn and predict microanatomical features based on different MRI contrasts. Similar to most projects that are heavily dependent on data, the availability of adequate and suitable datasets was one of the first challenges to tackle. For our project to work, we needed sets of images of different MRI contrasts and histology stains, ideally from the same animals. Even though medical datasets are nowadays in abundance and there are several individual MRI scans and atlases of different contrasts available online, we could not find - to the best of our knowledge - any publicly available sets of MRI and histology data of the same animals. Therefore, we had to acquire our own datasets from scratch. Since this was a secondary task to the project, the process of acquiring and processing the datasets was costly, time consuming, and required several iterations and revisions until satisfactory results were obtained. As a result, we were only able to acquire a full set of scans from one rat, including ex-vivo and in-vivo acquisitions with different contrasts and histology scans with different stains. Since we are considering 2D slices for the learning process, we had a sufficient number of coronal slices from MRI and histology scans

to kickstart our pipeline. We used patches and image augmentation (rotation and flipping) to overcome the shortage of data and expand our dataset size. However, since the data used to build the predictive models came from a single animal, it was not possible to validate whether the model has accurately learned the local histological features and would be able to map them from a new MRI scan of a different animal with biological variations. We therefore adopted an alternative technique to generally test if the neural network was indeed learning the desired features through predicting horizontal ex-vivo MRI slices from in-vivo slices using a model that was only trained on coronal slices. Even though the model was not able to predict all features from one anatomical view to another, the results were satisfactory, and some key features were successfully translated from the in-vivo image to the ex-vivo image. This shows that the model is in fact able to generalize to a certain degree and adapt to the variability introduced in the transition from a coronal view to a horizontal view.

5.1.2 Impact of registration quality

Another challenge that we had to work through was accurately registering the MRI images to their corresponding histology slices. There are several types of unwanted distortions that occur in both MRI and histology acquisitions, some of which can be addressed or accounted for, and others that cannot be avoided or fixed. For instance, MRI is subject to issues related to hardware (B0 and B1 inhomogeneities), software (errors in programming or shortcomings in the protocol), specimen (blood flow or motion) or physics (Gibbs ringing or susceptibility artifacts) [158]. Histology data is also subject to many potential distortions, mainly stemming from errors in sectioning leading to tears, folding artifacts or staining artifacts. These faults in the images will necessarily interfere with the

registration procedure and would therefore result in poorly registered pairs of MRI and histology features that will mislead the learning algorithm. Therefore, we had to optimize our registration strategy to maximize the quality of the input to the deep learning framework. We implemented the pipeline developed by Pilgram and Shmuel [137] with a few modifications. We used mini-blocks of histology slices to coarsely register to the 3D MRI volume and accurately extract the corresponding 2D MRI slice, with a precise z-location and a rotation angle. This was followed by a second stage of fine affine registration then finally a patch-based registration to optimize the localization of the 2D MRI slice that best corresponds the histology slice and registers well to it. Finally, we added an additional registration step of nonlinearly registering the patches that were being fed to the neural network and assessing their registration quality using the combined metric of SSIM and ECC to eliminate patches with poor registration. These registration improvements and data cleaning procedures guaranteed that all features being fed to the deep learning algorithm represented anatomically accurate information of the underlying structures, and therefore reduced the potential learning errors from distorted images or inaccurate registration.

Even though registration could be further improved using non-linear transformations to try and almost perfectly align MRI and histology voxel-by-voxel, it would not be the best practice. This is because, as mentioned earlier, the corresponding slices from MRI and histology are not identical in terms of underlying structures and trying to perfectly register them voxel-by-voxel would necessarily introduce deformations in key structures. In addition, we want the generated deep learning model to be able to generalize to different and new variations of the data without overfitting – which can be achieved better if there are

a few nonsignificant or natural misalignments between the images to help train the model on these variations.

5.1.3 Patch size optimization

The tradeoff faced while selecting a suitable patch size was subject to consideration and led to continuous tuning and optimization. Smaller patch sizes (approximately <20% of a coronal slice size) results in a more accurate mapping of the local features and is less influenced by large irrelevant structural features such as the external edges or shape of the brain slice. However, very small patches (<5% of the brain size) may also be misleading and result in very poor learning, since the registration between the two modalities of interest (MRI and histology) is not perfect, and features are not expected to align perfectly with one another at a pixel level. Therefore, we utilized skip connections between the encoder and the decoder in our implementation of the generator's UNet to allow the transfer of both high-level and low-level details directly through the net. In addition, we used multiple magnifications of patches in our training set. Meaning, the network was trained on patches that represented different structures at different magnifications to ensure that the model was learning the local features despite of some general morphological features that were repetitive along slices (such as the slice's outer edges or the shape of the corpus callosum). We used 3 different magnification level in our training set, representing approximately 10%, 15%, and 20% of the central coronal slice and applied to all other slices. Using image augmentation techniques (rotation and flipping) also aided our objective of mapping local features regardless of the structural appearance.

Moreover, the difference in resolution between histology and MRI adds an additional constraint to this optimization problem. If the modalities of interest were both of similar high resolutions, the size of the patches could decrease further and still be able to highlight meaningful structure and texture, which will therefore improve the learning of local features and map them more accurately. In our case, if the patch size is lower than <5% of the brain size, the ex-vivo MRI will represent nothing but noise. For in-vivo MRI, the resolution is even worse and the patch size cannot go lower than 10% of the brain size. This problem of resolution is believed to be one of the main constraints and main areas of improvement for this project.

5.2 Observations and analyses

Analyzing the predictions that resulted from training all the different models has allowed us to make a few observations. First, we noticed that almost all models have chosen not to map the folding artefacts in the histology during the learning process. For instance, it can be seen clearly in Figure 20 that even though the histology ground truth has a large folding artefact, this artefact was eliminated from the predicted histology and was not part of the learned features. This makes sense and aligns quite well with our expectations, since these artefacts do not appear on MRI images and therefore, it is likely that they have been outcasted by the neural network as an “outlier”. Moreover, this re-emphasizes the key concept that was discussed earlier about the rationale behind complementing MRI with histology, since histology acquisitions typically suffer from distortions and artefacts resulting from cutting, staining, and physical ex-vivo factors. These artefacts could be reduced or eliminated if histology was learned and predicted from in-vivo MRI images through a reliable pipeline.

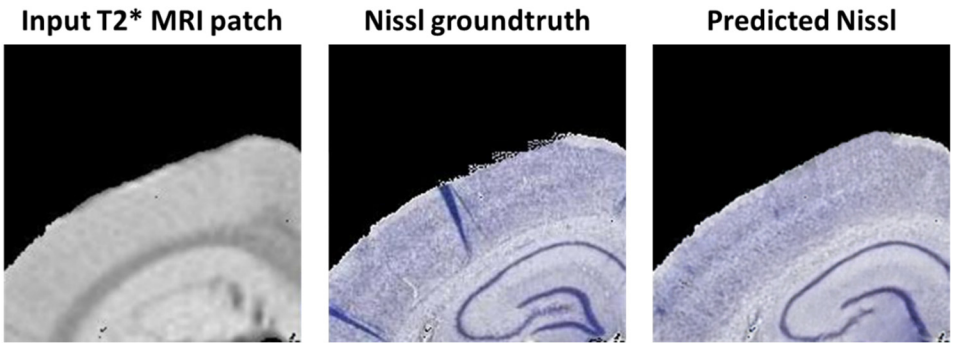

Figure 20: Elimination of the folding artifact by the trained model

The literature does not mention any direct relationships or correlations between Nissl and MRI parameters. Therefore, it is not expected that MRI contrasts would generally be good predictors of Nissl. We included Nissl histology slices in our analysis to validate our hypothesis and present it in contrast to other histological labels. As seen in Figure 14, Figure 16, and Figure 18 the models are able to predict Nissl but not as accurately as they predict myelin. The predictions of Nissl have less distinct features. This can also be confirmed by comparing the quantitative evaluation of Nissl against Myelin presented in Table 7, Table 8, and Table 9, where the quality of Nissl predictions (reflected by PSNR and SSIM metrics) is generally worse than the Myelin's predictions. Even though the current performance of the pipeline does not provide reliable quantitative predictions and merely serves as a proof of concept, this observation demonstrates the possibility of learning local features between MRI and histology and indicates that quantitative predictions can be achieved with improved implementations of this pipeline.

Another interesting observation that can be inferred from the qualitative and quantitative evaluation of the results is that there is a general trend of improved predictions with increasing information (i.e., the number of contrasts included in the learning). Even

though all experiments were implemented on the same data sample size, we notice that bi-modal predictions are usually better than uni-modal predictions, and multi-modal predictions are also generally better than bi-modal predictions. The way multi-modal and bi-modal experiments were handled is that MRI contrasts was fed as a multi-channel input, and the different contrasts were considered input feature maps. To ensure that the final sample size was the same across all experiments and to establish a fair comparison, we unified the number of feature channels across all experiments starting from the second layer. Meaning, if we started with a 1-channel input for uni-modal experiments, 2-channel input for bi-modal experiments, and a 3-channel input for multi-modal experiments, the second layer of all neural networks for all experiments will include 64 feature maps. This positive correlation between prediction quality and number of modalities can be explained by the variation of features covered by different MRI modalities and appearing on the histological slice. For instance, T2* images best highlight vascular detail, while the corpus collosum is most apparent on T1 images, and the ventricles are more defined on T2 images. The integration of these features together helps the neural network build more accurate patterns and perform a better mapping of the features.

5.3 Future work and directions

As highlighted earlier, qMRI measurements are anticipated to be the best predictors of microanatomical features and have been included in previous studies that investigated the possibility of achieving “in-vivo histology” [87]. Initially, we collected our MRI data with the intention to process them for quantitative data. We applied all necessary equations and tools to do so, as described in the literature review and methodology chapters. However, as mentioned earlier, since data acquisition was a challenging procedure, the fitted quantitative

measurements were too noisy to include in our pipeline. Also, since this study is intended to be a proof of concept in the first place, we decided to leave qMRI processing and analysis for future work and focus on analyzing MRI datasets of weighted measurements. Because we acquired data for qMRI, we had several images for T1-weighted contrast (with various TRs), T2-weighted contrast (with various TEs) and T2* (with various TEs) with multiple parameters, we only selected individual images that provided the best contrast and least noise. Finally, since we have already established the pipeline for qMRI processing, it will be easier to incorporate them in future studies with improved data acquisition.

Another interesting set of experiments to be investigated in the future is the ability of histological features to predict different MRI contrast. With sufficient data and appropriate implementation, these experiments could give important insights and conclusions regarding the contribution of microstructural features in shaping the MR signal and contrast. More broader studies could include additional MRI maps that might better reflect key microstructural features as described in the literature review. As established earlier, obtaining sufficient datasets with adequate resolution will have a great impact on the reliability findings.

Future work should also investigate the applicability of different machine learning and deep learning algorithms in building reliable mappings from MRI to histology and vice versa. Even though GANs are viewed as the state of the art for such image translation problems, different implementations could be of use under certain conditions. In the case that MRI and histology acquisition was optimized to the degree that corresponding slices could be aligned to one another in high confidence, polynomial regression could be used to

translate features voxel-by-voxel. For similar projects, the choice of the most suitable learning algorithm depends on the format of data available and is an open area of research that could drive interesting discussions and findings.

5.4 Conclusion

Through this work, we managed to fulfill the key objectives that were set in Chapter I of this manuscript. Namely, we:

1) Studied how microanatomical features presented in histological images can relate to their corresponding MRI images.
2) Developed an integrated pipeline with preprocessing, registration, and deep learning blocks to predict microanatomical features from MRI images without resorting to post-mortem analysis or invasive procedures.
3) Validated and tested our pipeline using through running several experiments to explore the relationship between different microstructural features (Nissl and Myelin) and MRI contrasts (T1w, T2w, and T2*w).

References

[1] N. Weiskopf, L. J. Edwards, G. Helms, S. Mohammadi, and E. Kirilina, "Quantitative magnetic resonance imaging of brain anatomy and in vivo histology," *Nat. Rev. Phys.*, vol. 3, no. 8, pp. 570–588, 2021.

[2] C. Hutton, B. Draganski, J. Ashburner, and N. Weiskopf, "A comparison between voxel-based cortical thickness and voxel-based morphometry in normal aging," *Neuroimage*, vol. 48, no. 2, pp. 371–380, 2009.

[3] J. Ashburner, J. Csernansky, C. Davatzikos, N. C. Fox, G. B. Frisoni, and P. M. Thompson, "Computer-assisted imaging to assess brain structure in healthy and diseased brains," *Lancet Neurol.*, vol. 2, pp. 79–88, 2003.

[4] J. Ashburner and K. J. Friston, "Voxel-based morphometry - The methods," *Neuroimage*, vol. 11, no. 6 I, pp. 805–821, 2000.

[5] B. Fischl and A. M. Dale, "Measuring the thickness of the human cerebral cortex from magnetic resonance images," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 97, no. 20, pp. 11050–11055, 2000.

[6] M. F. Callaghan *et al.*, "Widespread age-related differences in the human brain microstructure revealed by quantitative magnetic resonance imaging," *Neurobiol. Aging*, vol. 35, no. 8, pp. 1862–1872, 2014.

[7] C. D. Good, I. S. Johnsrude, J. Ashburner, R. N. A. Henson, K. J. Friston, and R. S. J. Frackowiak, "A voxel-based morphometric study of ageing in 465 normal adult human brains," *Neuroimage*, vol. 14, no. 1 I, pp. 21–36, 2001.

[8] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens, and P. M. Thompson, "The clinical use of structural MRI in Alzheimer disease," *Nat. Rev. Neurol.*, vol. 6, no. 2, pp. 67–77, 2010.

[9] R. J. Zatorre, R. D. Fields, and H. Johansen-Berg, "Plasticity in gray and white: Neuroimaging changes in brain structure during learning," *Nat. Neurosci.*, vol. 15, no. 4, pp. 528–536, Apr. 2012.

[10] A. M. Maillard *et al.*, "The 16p11.2 locus modulates brain structures common to autism, schizophrenia and obesity," *Mol. Psychiatry*, vol. 20, no. 1, pp. 140–147, 2015.

[11] E. M. Shapiro, K. Sharer, S. Skrtic, and A. P. Koretsky, "In vivo detection of single cells by MRI," *Magn. Reson. Med.*, vol. 55, no. 2, pp. 242–249, Feb. 2006.

[12] Y. Assaf, "Imaging laminar structures in the gray matter with diffusion MRI," *Neuroimage*, vol. 197, pp. 677–688, 2019.

[13] M. D. Does, "Inferring brain tissue composition and microstructure via MR relaxometry," *Neuroimage*, vol. 182, pp. 136–148, Nov. 2018.

[14] T. E. Yankeelov, R. G. Abramson, and C. C. Quarles, "Quantitative multimodality imaging in cancer research and therapy," *Nat. Rev. Clin. Oncol.*, vol. 11, no. 11, pp. 670–680, Nov. 2014.

[15] M. Incoronato *et al.*, "Radiogenomic Analysis of Oncological Data: A Technical Survey," *Int. J. Mol. Sci.*, vol. 18, no. 4, p. 805, Apr. 2017.

[16] R. G. Abramson *et al.*, "MR Imaging Biomarkers in Oncology Clinical Trials," *Magn. Reson.*

Imaging Clin. N. Am., vol. 24, no. 1, pp. 11–29, Feb. 2016.

[17] P. Duvnjak, A. A. Schulman, J. N. Holtz, J. Huang, T. J. Polascik, and R. T. Gupta, "Multiparametric Prostate MR Imaging: Impact on Clinical Staging and Decision Making," *Radiol. Clin.*, vol. 56, no. 2, pp. 239–250, Mar. 2018.

[18] K. Lovett, M. D. Rifkin, P. A. McCue, and H. Choi, "MR imaging characteristics of noncancerous lesions of the prostate," *J. Magn. Reson. Imaging*, vol. 2, no. 1, pp. 35–39, Jan. 1992.

[19] M. K. Brawer, R. E. Deering, M. Brown, S. D. Preston, and S. A. Bigler, "Predictors of pathologic stage in prostatic carcinoma. The role of neovascularity," *Cancer*, vol. 73, no. 3, pp. 678–687, Feb. 1994.

[20] R. M. Bourne *et al.*, "Apparatus for Histological Validation of In Vivo and Ex Vivo Magnetic Resonance Imaging of the Human Prostate," *Front. Oncol.*, vol. 7, p. 47, Mar. 2017.

[21] S. V. Ramagopalan, R. Dobson, U. C. Meier, and G. Giovannoni, "Multiple sclerosis: risk factors, prodromes, and potential causal pathways," *Lancet Neurol.*, vol. 9, no. 7, pp. 727–739, 2010.

[22] A. Alizadeh, S. M. Dyck, and S. Karimi-Abdolrezaee, "Myelin damage and repair in pathologic CNS: Challenges and prospects," *Front. Mol. Neurosci.*, vol. 8, no. JULY, pp. 1–27, 2015.

[23] F. Heath, S. A. Hurley, H. Johansen-Berg, and C. Sampaio-Baptista, "Advances in noninvasive myelin imaging," *Dev. Neurobiol.*, vol. 78, no. 2, pp. 136–151, 2018.

[24] C. W. J. van der Weijden *et al.*, "Myelin quantification with MRI: A systematic review of accuracy and reproducibility," *Neuroimage*, vol. 226, no. November 2020, 2021.

[25] W. T. Blume, G. R. Ganapathy, D. Munoz, and D. H. Lee, "Indices of Resective Surgery Effectiveness for Intractable Nonlesional Focal Epilepsy," *Epilepsia*, vol. 45, no. 1, pp. 46–53, 2004.

[26] L. E. Jeha *et al.*, "Predictors of outcome after temporal lobectomy for the treatment of intractable epilepsy," *Neurology*, vol. 66, no. 12, pp. 1938 LP – 1940, Jun. 2006.

[27] A. A. Cohen-Gadol *et al.*, "Normal magnetic resonance imaging and medial temporal lobe epilepsy: the clinical syndrome of paradoxical temporal lobe epilepsy," *J. Neurosurg.*, vol. 102, no. 5, pp. 902–909, 2005.

[28] P. N. Sylaja, K. Radhakrishnan, C. Kesavadas, and P. S. Sarma, "Seizure Outcome after Anterior Temporal Lobectomy and Its Predictors in Patients with Apparent Temporal Lobe Epilepsy and Normal MRI," *Epilepsia*, vol. 45, no. 7, pp. 803–808, Jul. 2004.

[29] A. Bernasconi *et al.*, "T2 Relaxometry Can Lateralize Mesial Temporal Lobe Epilepsy in Patients with Normal MRI," *Neuroimage*, vol. 12, no. 6, pp. 739–746, 2000.

[30] N. Bernasconi, S. Duchesne, A. Janke, J. Lerch, D. L. Collins, and A. Bernasconi, "Whole-brain voxel-based statistical analysis of gray matter and white matter in temporal lobe epilepsy," *Neuroimage*, vol. 23, no. 2, pp. 717–723, 2004.

[31] B. C. Bernhardt, K. J. Worsley, H. Kim, A. C. Evans, A. Bernasconi, and N. Bernasconi, "Longitudinal and cross-sectional analysis of atrophy in pharmacoresistant temporal lobe epilepsy," *Neurology*, vol. 72, no. 20, pp. 1747–1754, May 2009.

[32] R. Garbelli *et al.*, "Blurring in patients with temporal lobe epilepsy: clinical, high-field imaging and ultrastructural study," *Brain*, vol. 135, no. 8, pp. 2337–2349, Aug. 2012.

[33] S. H. Eriksson *et al.*, "Correlation of quantitative MRI and neuropathology in epilepsy surgical resection specimens--T2 correlates with neuronal tissue in gray matter," *Neuroimage*, vol. 37, no. 1, pp. 48–55, Aug. 2007.

[34] S. H. Eriksson *et al.*, "Quantitative grey matter histological measures do not correlate with grey matter probability values from in vivo MRI in the temporal lobe," *J. Neurosci. Methods*, vol. 181, no. 1, pp. 111–118, Jun. 2009.

[35] G. Lockwood-Estrin *et al.*, "Correlating 3T MRI and histopathology in patients undergoing epilepsy surgery," *J. Neurosci. Methods*, vol. 205, no. 1, pp. 182–189, 2012.

[36] M. Goubran *et al.*, "Magnetic resonance imaging and histology correlation in the neocortex in temporal lobe epilepsy," *Ann. Neurol.*, vol. 77, no. 2, pp. 237–250, 2015.

[37] O. Jahn, S. Tenzer, and H. B. Werner, "Myelin proteomics: molecular anatomy of an insulating sheath," *Mol. Neurobiol.*, vol. 40, no. 1, pp. 55–72, 2009.

[38] C. Laule, P. Kozlowski, E. Leung, D. K. B. Li, A. L. MacKay, and G. R. W. Moore, "Myelin water imaging of multiple sclerosis at 7 T: correlations with histopathology," *Neuroimage*, vol. 40, no. 4, pp. 1575–1580, 2008.

[39] B. Drayer, P. Burger, R. Darwin, S. Riederer, R. Herfkens, and G. A. Johnson, "MRI of brain iron," *Am. J. Roentgenol.*, vol. 147, no. 1, pp. 103–110, 1986.

[40] S. Geyer, M. Weiss, K. Reimann, G. Lohmann, and R. Turner, "Microstructural parcellation of the human cerebral cortex–from Brodmann's post-mortem map to in vivo mapping with high-field magnetic resonance imaging," *Front. Hum. Neurosci.*, vol. 5, p. 19, 2011.

[41] I. S. Sigalovsky, B. Fischl, and J. R. Melcher, "Mapping an intrinsic MR property of gray matter in auditory cortex of living humans: a possible marker for primary cortex and hemispheric differences," *Neuroimage*, vol. 32, no. 4, pp. 1524–1537, 2006.

[42] M. F. Glasser and D. C. Van Essen, "Mapping human cortical areas in vivo based on myelin content as revealed by T1-and T2-weighted MRI," *J. Neurosci.*, vol. 31, no. 32, pp. 11597–11616, 2011.

[43] N. Gelman, J. R. Ewing, J. M. Gorell, E. M. Spickler, and E. G. Solomon, "Interregional variation of longitudinal relaxation rates in human brain at 3.0 T: Relation to estimated iron and water contents," *Magn. Reson. Med.*, vol. 45, no. 1, pp. 71–79, Jan. 2001.

[44] Y. Gossuin, A. Roch, R. N. Muller, and P. Gillis, "Relaxation induced by ferritin and ferritin-like magnetic particles: the role of proton exchange," *Magn. Reson. Med. An Off. J. Int. Soc. Magn. Reson. Med.*, vol. 43, no. 2, pp. 237–243, 2000.

[45] R. G. Steen, W. E. Reddick, and R. J. Ogg, "More than meets the eye: significant regional heterogeneity in human cortical T1☆," *Magn. Reson. Imaging*, vol. 18, no. 4, pp. 361–368, 2000.

[46] R. J. Ogg and R. G. Steen, "Age-related changes in brain T1 are correlated with iron concentration," *Magn. Reson. Med.*, vol. 40, no. 5, pp. 749–753, 1998.

[47] J. Vymazal *et al.*, "The quantitative relation between T1-weighted and T2-weighted MRI of normal gray matter and iron concentration.," *J. Magn. Reson. imaging JMRI*, vol. 5, no. 5, pp. 554–560, 1995.

[48] G. B. Chavhan, P. S. Babyn, B. Thomas, M. M. Shroff, and E. M. Haacke, “Principles, techniques, and applications of T2*-based MR imaging and its special applications,” *Radiographics*, vol. 29, no. 5, pp. 1433–1449, 2009.

[49] R. J. Ordidge, J. M. Gorell, J. C. Deniau, R. A. Knight, and J. A. Helpern, “Assessment of relative brain iron concentrations using T2-weighted and T2*-weighted MRI at 3 Tesla,” *Magn. Reson. Med.*, vol. 32, no. 3, pp. 335–341, 1994.

[50] J. F. Schenck and E. A. Zimmerman, “High-field magnetic resonance imaging of brain iron: birth of a biomarker?,” *NMR Biomed. An Int. J. Devoted to Dev. Appl. Magn. Reson. Vivo*, vol. 17, no. 7, pp. 433–445, 2004.

[51] A. Schäfer *et al.*, “Direct visualization of the subthalamic nucleus and its iron distribution using high-resolution susceptibility mapping,” *Hum. Brain Mapp.*, vol. 33, no. 12, pp. 2831–2842, 2012.

[52] J. P. Marques and R. Bowtell, “Application of a Fourier-based method for rapid calculation of field inhomogeneity due to spatial variation of magnetic susceptibility,” *Concepts Magn. Reson. Part B Magn. Reson. Eng.*, vol. 25B, no. 1, pp. 65–78, Apr. 2005.

[53] S. Wharton and R. Bowtell, “Fiber orientation-dependent white matter contrast in gradient echo MRI,” *Proc. Natl. Acad. Sci.*, vol. 109, no. 45, pp. 18559 LP – 18564, Nov. 2012.

[54] J. Duyn, “MR susceptibility imaging,” *J. Magn. Reson.*, vol. 229, pp. 198–207, 2013.

[55] X. He and D. A. Yablonskiy, “Biophysical mechanisms of phase contrast in gradient echo MRI,” *Proc. Natl. Acad. Sci.*, vol. 106, no. 32, pp. 13558 LP – 13563, Aug. 2009.

[56] S. C. L. Deoni, S. C. R. Williams, P. Jezzard, J. Suckling, D. G. M. Murphy, and D. K. Jones, “Standardized structural magnetic resonance imaging in multicentre studies using quantitative T1 and T2 imaging at 1.5 T,” *Neuroimage*, vol. 40, no. 2, pp. 662–671, 2008.

[57] G. Helms, H. Dathe, and P. Dechent, “Quantitative FLASH MRI at 3T using a rational approximation of the Ernst equation,” *Magn. Reson. Med.*, vol. 59, no. 3, pp. 667–672, Mar. 2008.

[58] C. Langkammer *et al.*, “Quantitative MR Imaging of Brain Iron: A Postmortem Validation Study,” *Radiology*, vol. 257, no. 2, pp. 455–462, Nov. 2010.

[59] J. Cohen-Adad, “What can we learn from T2* maps of the cortex?,” *Neuroimage*, vol. 93, pp. 189–200, 2014.

[60] H. Neeb, V. Ermer, T. Stocker, and N. J. Shah, “Fast quantitative mapping of absolute water content with full brain coverage,” *Neuroimage*, vol. 42, no. 3, pp. 1094–1109, 2008.

[61] N. Weiskopf *et al.*, “Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: a multi-center validation ,” *Frontiers in Neuroscience* , vol. 7. p. 95, 2013.

[62] C. Laule *et al.*, “Myelin water imaging in multiple sclerosis: quantitative correlations with histopathology,” *Mult. Scler. J.*, vol. 12, no. 6, pp. 747–753, Nov. 2006.

[63] A. M. Maillard *et al.*, “The clinical use of structural MRI in Alzheimer disease,” *Neuroimage*, vol. 6, no. 2, pp. 805–821, 2000.

[64] G. Helms, H. Dathe, K. Kallenberg, and P. Dechent, “High-resolution maps of magnetization transfer with inherent correction for RF inhomogeneity and T1 relaxation obtained from 3D

FLASH MRI," *Magn. Reson. Med.*, vol. 60, no. 6, pp. 1396–1407, Dec. 2008.

[65] D. K. Müller, A. Pampel, and H. E. Möller, "Matrix-algebra-based calculations of the time evolution of the binary spin-bath model for magnetization transfer," *J. Magn. Reson.*, vol. 230, pp. 88–97, 2013.

[66] S. Kalantari, C. Laule, T. A. Bjarnason, I. M. Vavasour, and A. L. MacKay, "Insight into in vivo magnetization exchange in human white matter regions," *Magn. Reson. Med.*, vol. 66, no. 4, pp. 1142–1151, Oct. 2011.

[67] K. Schmierer, F. Scaravilli, D. R. Altmann, G. J. Barker, and D. H. Miller, "Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain," *Ann. Neurol.*, vol. 56, no. 3, pp. 407–415, Sep. 2004.

[68] A. Lutti, F. Dick, M. I. Sereno, and N. Weiskopf, "Using high-resolution quantitative mapping of R1 as an index of cortical myelination," *Neuroimage*, vol. 93, pp. 176–188, 2014.

[69] H. Johansen-Berg and T. E. J. B. T.-D. M. R. I. (Second E. Behrens, Eds., "Index," San Diego: Academic Press, 2014, pp. 607–614.

[70] P. Jones Derek K., *Diffusion MRITheory, Methods, and Applications: Theory, Methods, and Applications*. Oxford, UK: Oxford University Press, 2012.

[71] Y. Assaf and P. J. Basser, "Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain," *Neuroimage*, vol. 27, no. 1, pp. 48–58, 2005.

[72] D. Barazany, P. J. Basser, and Y. Assaf, "In vivo measurement of axon diameter distribution in the corpus callosum of rat brain," *Brain*, vol. 132, no. Pt 5, pp. 1210–1220, May 2009.

[73] H. Zhang, T. Schneider, C. A. Wheeler-Kingshott, and D. C. Alexander, "NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain," *Neuroimage*, vol. 61, no. 4, pp. 1000–1016, 2012.

[74] C. Thomas *et al.*, "Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 46, pp. 16574–16579, Nov. 2014.

[75] M. Reisert, V. G. Kiselev, B. Dihtal, E. Kellner, and D. S. Novikov, "MesoFT: unifying diffusion modelling and fiber tracking," *Med. Image Comput. Comput. Assist. Interv.*, vol. 17, no. Pt 3, pp. 201–208, 2014.

[76] M. Kleinnijenhuis, T. van Mourik, D. G. Norris, D. J. Ruiter, A.-M. van Cappellen van Walsum, and M. Barth, "Diffusion tensor characteristics of gyrencephaly using high resolution diffusion MRI in vivo at 7T," *Neuroimage*, vol. 109, pp. 378–387, 2015.

[77] A. K. Seehaus *et al.*, "Histological validation of DW-MRI tractography in human postmortem tissue," *Cereb. Cortex*, vol. 23, no. 2, pp. 442–450, Feb. 2013.

[78] G. Zhao, D. K. B. Li, and D. Paty, "15 - MRI in Multiple Sclerosis," J. C. Mazziotta, A. W. Toga, and R. S. J. B. T.-B. M. T. D. Frackowiak, Eds. San Diego: Academic Press, 2000, pp. 357–385.

[79] P. S. Tofts, "The Measurement Process: MR Data Collection and Image Analysis," *Quantitative MRI of the Brain*. pp. 17–54, 15-Aug-2003.

[80] N. K. Focke *et al.*, "Multi-site voxel-based morphometry — Not quite there yet," *Neuroimage*, vol. 56, no. 3, pp. 1164–1170, 2011.

[81] G. Helms, B. Draganski, R. Frackowiak, J. Ashburner, and N. Weiskopf, “Improved segmentation of deep brain grey matter structures using magnetization transfer (MT) parameter maps,” *Neuroimage*, vol. 47, no. 1, pp. 194–198, Aug. 2009.

[82] G. Helms, H. Dathe, and P. Dechent, “Modeling the influence of TR and excitation flip angle on the magnetization transfer ratio (MTR) in human brain obtained from 3D spoiled gradient echo MRI,” *Magn. Reson. Med.*, vol. 64, no. 1, pp. 177–185, Jul. 2010.

[83] C. Lambert, A. Lutti, G. Helms, R. Frackowiak, and J. Ashburner, “Multiparametric brainstem segmentation using a modified multivariate mixture of Gaussians,” *NeuroImage Clin.*, vol. 2, pp. 684–694, 2013.

[84] S. Lorio *et al.*, “Disentangling in vivo the effects of iron content and atrophy on the ageing human brain,” *Neuroimage*, vol. 103, pp. 280–289, Dec. 2014.

[85] S. C. L. Deoni, T. M. Peters, and B. K. Rutt, “High-resolution T1 and T2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2,” *Magn. Reson. Med.*, vol. 53, no. 1, pp. 237–241, Jan. 2005.

[86] A. Mezer *et al.*, “Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging,” *Nat. Med.*, vol. 19, no. 12, pp. 1667–1672, Dec. 2013.

[87] N. Weiskopf, S. Mohammadi, A. Lutti, and M. F. Callaghan, “Advances in MRI-based computational neuroanatomy: From morphometry to in-vivo histology,” *Curr. Opin. Neurol.*, vol. 28, no. 4, pp. 313–322, 2015.

[88] K. Tabelow *et al.*, “hMRI–A toolbox for quantitative MRI in neuroscience and clinical research,” *Neuroimage*, vol. 194, pp. 191–210, 2019.

[89] K. Friston, “A short history of SPM,” *Stat. Parametr. Mapp. Anal. Funct. brain images*, pp. 3–9, 2007.

[90] A. Karakuzu *et al.*, “qMRLab: Quantitative MRI analysis, under one umbrella,” *J. Open Source Softw.*, vol. 5, no. 53, p. 2343, 2020.

[91] J. A. McNab *et al.*, “The Human Connectome Project and beyond: Initial applications of 300mT/m gradients,” *Neuroimage*, vol. 80, pp. 234–245, 2013.

[92] K. Ugurbil, “Magnetic resonance imaging at ultrahigh fields,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1364–1379, May 2014.

[93] M. F. Callaghan, O. Josephs, M. Herbst, M. Zaitsev, N. Todd, and N. Weiskopf, “An evaluation of prospective motion correction (PMC) for high resolution quantitative MRI,” *Front. Neurosci.*, vol. 9, p. 97, Mar. 2015.

[94] N. Weiskopf, M. F. Callaghan, O. Josephs, A. Lutti, and S. Mohammadi, “Estimating the apparent transverse relaxation time (R2(*)) from images with different contrasts (ESTATICS) reduces motion artifacts,” *Front. Neurosci.*, vol. 8, p. 278, Sep. 2014.

[95] S. Mohammadi, C. Hutton, Z. Nagy, O. Josephs, and N. Weiskopf, “Retrospective correction of physiological noise in DTI using an extended tensor model and peripheral measurements,” *Magn. Reson. Med.*, vol. 70, no. 2, pp. 358–369, Aug. 2013.

[96] F. Calamante *et al.*, “Super-resolution track-density imaging studies of mouse brain: Comparison to histology,” *Neuroimage*, vol. 59, no. 1, pp. 286–296, 2012.

[97] L. Ruthotto, S. Mohammadi, and N. Weiskopf, “A new method for joint susceptibility artefact correction and super-resolution for dMRI,” in *Proc.SPIE*, 2014, vol. 9034.

[98] S. M. A. Becker, K. Tabelow, S. Mohammadi, N. Weiskopf, and J. Polzehl, “Adaptive smoothing of multi-shell diffusion weighted magnetic resonance data by msPOAS,” *Neuroimage*, vol. 95, pp. 90–105, 2014.

[99] C. A. Wheeler-Kingshott *et al.*, “The current state-of-the-art of spinal cord imaging: applications,” *Neuroimage*, vol. 84, pp. 1082–1093, Jan. 2014.

[100] M. Deppe *et al.*, “Diffusion-Tensor Imaging at 3 T: Detection of White Matter Alterations in Neurological Patients on the Basis of Normal Values,” *Invest. Radiol.*, vol. 42, no. 6, 2007.

[101] A. Deistung *et al.*, “Quantitative susceptibility mapping differentiates between blood depositions and calcifications in patients with glioblastoma,” *PLoS One*, vol. 8, no. 3, pp. e57924–e57924, 2013.

[102] M. F. Callaghan, G. Helms, A. Lutti, S. Mohammadi, and N. Weiskopf, “A general linear relaxometry model of R1 using imaging data,” *Magn. Reson. Med.*, vol. 73, no. 3, pp. 1309–1314, Mar. 2015.

[103] C. Stüber *et al.*, “Myelin and iron concentration in the human brain: A quantitative study of MRI contrast,” *Neuroimage*, vol. 93, pp. 95–106, 2014.

[104] A. MacKay, C. Laule, I. Vavasour, T. Bjarnason, S. Kolind, and B. Mädler, “Insights into brain microstructure from the T2 distribution,” *Magn. Reson. Imaging*, vol. 24, no. 4, pp. 515–525, 2006.

[105] S. N. Jespersen *et al.*, “Neurite density from magnetic resonance diffusion measurements at ultrahigh field: Comparison with light microscopy and electron microscopy,” *Neuroimage*, vol. 49, no. 1, pp. 205–216, 2010.

[106] W. W. Lam, S. Jbabdi, and K. L. Miller, “A model for extra-axonal diffusion spectra with frequency-dependent restriction,” *Magn. Reson. Med.*, vol. 73, no. 6, pp. 2306–2320, Jun. 2015.

[107] M. A. Koch and J. Finsterbusch, “Compartment size estimation with double wave vector diffusion-weighted imaging,” *Magn. Reson. Med.*, vol. 60, no. 1, pp. 90–101, Jul. 2008.

[108] D. A. Feinberg *et al.*, “Multiplexed Echo Planar Imaging for Sub-Second Whole Brain FMRI and Fast Diffusion Imaging,” *PLoS One*, vol. 5, no. 12, p. e15710, Dec. 2010.

[109] S. N. Sotiropoulos *et al.*, “Advances in diffusion MRI acquisition and processing in the Human Connectome Project,” *Neuroimage*, vol. 80, pp. 125–143, 2013.

[110] J. R. Reichenbach, “The future of susceptibility contrast for assessment of anatomy and function,” *Neuroimage*, vol. 62, no. 2, pp. 1311–1315, 2012.

[111] B. Todorich, J. M. Pasquini, C. I. Garcia, P. M. Paez, and J. R. Connor, “Oligodendrocytes and myelination: The role of iron,” *Glia*, vol. 57, no. 5, pp. 467–478, Apr. 2009.

[112] W. C. Chen, S. Foxley, and K. L. Miller, “Detecting microstructural properties of white matter based on compartmentalization of magnetic susceptibility,” *Neuroimage*, vol. 70, pp. 1–9, 2013.

[113] J. H. Duyn, “Frequency shifts in the myelin water compartment,” *Magn. Reson. Med.*, vol. 71,

no. 6, pp. 1953–1955, Jun. 2014.

[114] P. Sati *et al.*, "Micro-compartment specific T2* relaxation in the brain," *Neuroimage*, vol. 77, pp. 268–278, Aug. 2013.

[115] F. P. M. Oliveira and J. M. R. S. Tavares, "Medical image registration: a review," *Comput. Methods Biomech. Biomed. Engin.*, vol. 17, no. 2, pp. 73–93, Jan. 2014.

[116] J. V. Hajnal and D. L. G. Hill, *Medical Image Registration*, 1st Editio. CRC Press, 2001.

[117] S. K. Boopathy Jegathambal, K. Mok, D. A. Rudko, and A. Shmuel, "MRI Based Brain-Specific 3D-Printed Model Aligned to Stereotactic Space for Registering Histology to MRI.," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2018, pp. 802–805, Jul. 2018.

[118] H. Livyatan, Z. Yaniv, and L. Joskowicz, "Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT," *IEEE Trans. Med. Imaging*, vol. 22, no. 11, pp. 1395–1406, 2003.

[119] M. Auer, P. Regitnig, and G. A. Holzapfel, "An automatic nonrigid registration for stained histological sections," *IEEE Trans. Image Process.*, vol. 14, no. 4, pp. 475–486, 2005.

[120] N. J. AU - Luciano *et al.*, "Utilizing 3D Printing Technology to Merge MRI with Histology: A Protocol for Brain Sectioning," *JoVE*, no. 118, p. e54780, 2016.

[121] A. Jog, A. Carass, S. Roy, D. L. Pham, and J. L. Prince, "Random forest regression for magnetic resonance image synthesis," *Med. Image Anal.*, vol. 35, pp. 475–488, 2017.

[122] M. I. Miller, G. E. Christensen, Y. Amit, and U. Grenander, "Mathematical textbook of deformable neuroanatomies," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 90, no. 24, pp. 11944–11948, Dec. 1993.

[123] A. Jog, A. Carass, S. Roy, D. L. Pham, and J. L. Prince, "MR image synthesis by contrast learning on neighborhood ensembles," *Med. Image Anal.*, vol. 24, no. 1, pp. 63–76, Aug. 2015.

[124] A. Jog, S. Roy, A. Carass, and J. L. Prince, "Magnetic resonance image synthesis through patch regression," in *2013 IEEE 10th International Symposium on Biomedical Imaging*, 2013, pp. 350–353.

[125] S. Roy, A. Carass, and J. Prince, "A compressed sensing approach for MR tissue contrast synthesis," *Inf. Process. Med. Imaging*, vol. 22, pp. 371–383, 2011.

[126] S. Roy, A. Jog, A. Carass, and J. L. Prince, "Atlas Based Intensity Transformation of Brain MR Images," in *Proceedings of Third International Workshop on Multimodal Brain Image Analysis - Volume 8159*, 2013, pp. 51–62.

[127] N. Cordier, H. Delingette, M. Lê, and N. Ayache, "Extended modality propagation: image synthesis of pathological cases," *IEEE Trans. Med. Imaging*, vol. 35, no. 12, pp. 2598–2608, 2016.

[128] D. H. Ye, D. Zikic, B. Glocker, A. Criminisi, and E. Konukoglu, "Modality propagation: coherent synthesis of subject-specific scans with data-driven regularization," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2013, pp. 606–613.

[129] H. Van Nguyen, K. Zhou, and R. Vemulapalli, "Cross-domain synthesis of medical images using efficient location-sensitive deep network," in *International Conference on Medical*

Image Computing and Computer-Assisted Intervention, 2015, pp. 677–684.

[130] A. Chartsias, T. Joyce, M. V. Giuffrida, and S. A. Tsaftaris, "Multimodal MR synthesis via modality-invariant latent representation," *IEEE Trans. Med. Imaging*, vol. 37, no. 3, pp. 803–814, 2017.

[131] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *European conference on computer vision*, 2016, pp. 694–711.

[132] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[133] A. Dosovitskiy and T. Brox, "Generating images with perceptual similarity metrics based on deep networks," *Adv. Neural Inf. Process. Syst.*, vol. 29, pp. 658–666, 2016.

[134] S. Bauer, T. Fejes, and M. Reyes, "A Skull-Stripping Filter for ITK." 2013.

[135] J. S. Lim, "Two-dimensional signal and image processing," *Englewood Cliffs*, 1990.

[136] R. C. Gonzalez, R. E. Woods, and B. R. Masters, "Digital image processing." Society of Photo-Optical Instrumentation Engineers, 2009.

[137] R. Pilgram, "Development of methods and applications to the study of the structure, function, and functional connectivity of cortical layers," McGill University, 2021.

[138] J.-P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, 1998.

[139] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "PET-CT image registration in the chest using free-form deformations," *IEEE Trans. Med. Imaging*, vol. 22, no. 1, pp. 120–128, 2003.

[140] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[141] J. Astola and I. Virtanen, *Entropy correlation coefficient, a measure of statistical dependence for categorized data*. Lappeenrannan teknillinen korkeakoulu, 1981.

[142] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognit.*, vol. 32, no. 1, pp. 71–86, 1999.

[143] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "Elastix: a toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imaging*, vol. 29, no. 1, pp. 196–205, 2009.

[144] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv Prepr. arXiv1411.1784*, 2014.

[145] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 5967–5976, 2017.

[146] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE conference on computer vision and pattern*

recognition, 2016, pp. 2536–2544.

[147] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015*, 2015, pp. 234–241.

[148] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther, "Autoencoding beyond pixels using a learned similarity metric," in *International conference on machine learning*, 2016, pp. 1558–1566.

[149] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proceedings of the seventh IEEE international conference on computer vision*, 1999, vol. 2, pp. 1033–1038.

[150] L. Gatys, A. S. Ecker, and M. Bethge, "Texture synthesis using convolutional neural networks," *Adv. Neural Inf. Process. Syst.*, vol. 28, 2015.

[151] C. Li and M. Wand, "Precomputed real-time texture synthesis with markovian generative adversarial networks," in *European conference on computer vision*, 2016, pp. 702–716.

[152] I. Goodfellow *et al.*, "Generative adversarial nets," *Adv. Neural Inf. Process. Syst.*, vol. 27, 2014.

[153] J. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

[154] B. Draganski *et al.*, "Regional specificity of MRI contrast parameter changes in normal ageing revealed by voxel-based quantification (VBQ)," *Neuroimage*, vol. 55, no. 4, pp. 1423–1434, 2011.

[155] P. Freund *et al.*, "MRI investigation of the sensorimotor cortex and the corticospinal tract after acute spinal cord injury: a prospective longitudinal study," *Lancet Neurol.*, vol. 12, no. 9, pp. 873–881, 2013.

[156] G. Iannucci, M. Rovaris, L. Giacomotti, G. Comi, and M. Filippi, "Correlation of multiple sclerosis measures derived from T2-weighted, T1-weighted, magnetization transfer, and diffusion tensor MR imaging," *Am. J. Neuroradiol.*, vol. 22, no. 8, pp. 1462–1467, 2001.

[157] S. Helbling *et al.*, "Structure predicts function: Combining non-invasive electrophysiology with in-vivo histology," *Neuroimage*, vol. 108, pp. 377–385, 2015.

[158] K. Somasundaram and P. Kalavathi, "Analysis of Imaging Artifacts in MR Brain Images," *Orient. J. Comput. Sci. …*, vol. 5 N0 1, no. June 2012, pp. 135–141, 2012.

Appendix I: Abbreviations

MRI – Magnetic Resonance Imaging

qMRI – Quantitative Magnetic Resonance Imaging

NMR – Nuclear Magnetic Resonance

CT – Computed Tomography

SPECT – Single-Photon Emission Computed Tomography

MS – Multiple Sclerosis

CNS – Central Nervous System

DTI – Diffusion Tensor Imaging

MWF – Myelin Water Fraction

FLAIR - Fast Fluid-Attenuated Inversion Recovery

NeuN - Neuronal Nuclear Antigen

QSM – Quantitative Susceptibility Mapping

RF – Radiofrequency

VBM – Voxel-based Morphometry

PD – Proton Density

MT – Magnetization Transfer

SPM – Statistical Parametric Mapping

GUI – Graphical User Interface

hMRI – In-vivo Histology

DWI – Diffusion Weighted Imaging

DOF – Degrees of Freedom

MI – Mutual Information

MSE – Mean Squared Error

GAN – Generative Adversarial Network

TR – Repetition Time

TE – Echo Time

ME-SE – Multi-echo Spin-echo

ME-GRE – Multi-echo Gradient-echo

VTR – Variable Repetition Time

SNR – Signal to Noise Ratio

FOV – Field of View

FA – Flip Angle

IR – Inversion Recovery

RARE – Relaxation Enhancement

ETL – Echo Train Length

SSIM – Structural Similarity Index

ECC – Entropy Correlation Coefficient

Appendix II: Additional examples on registration evaluation

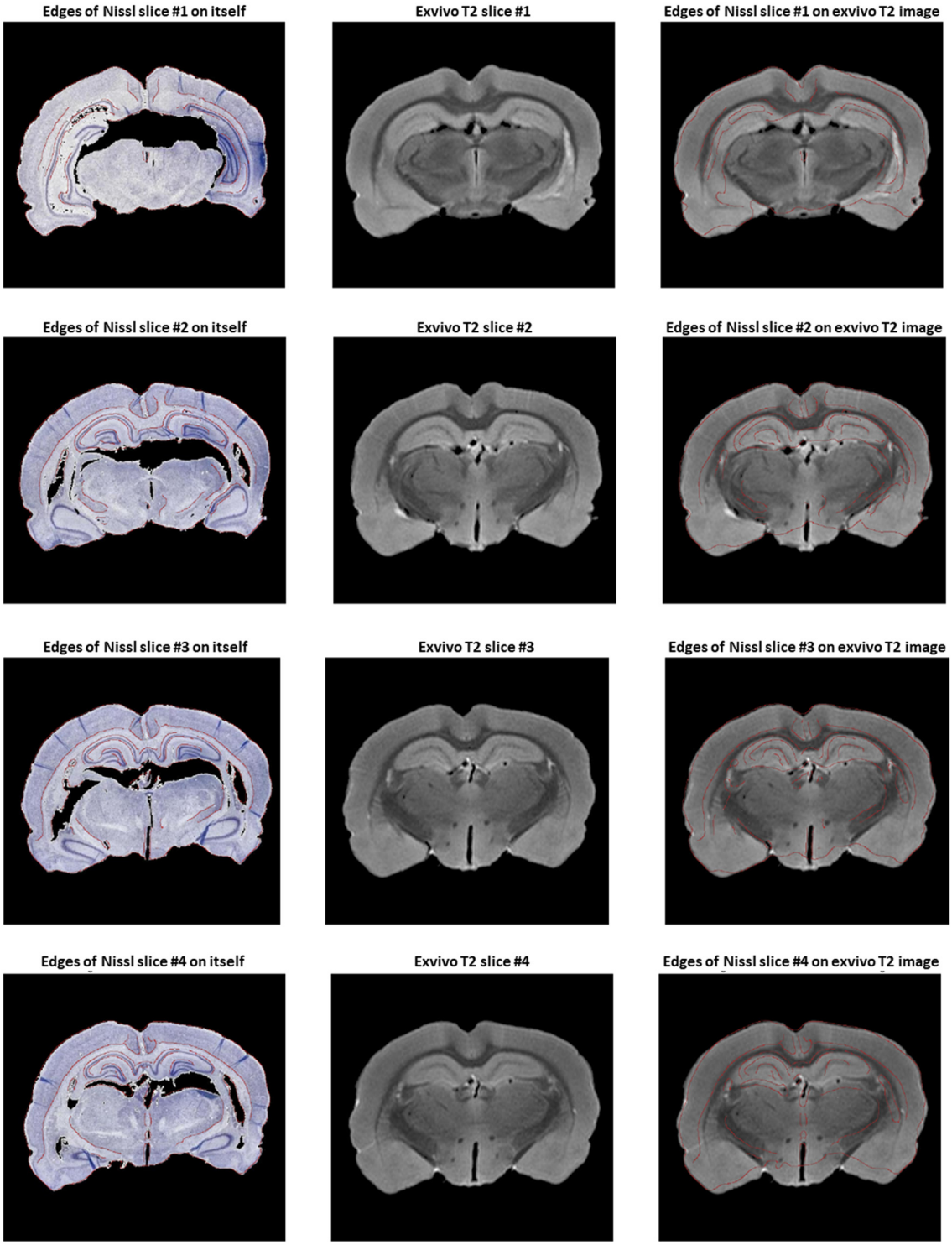

Figure 21: Qualitative evaluation using superimposed edges of the registration between T2 MRI and corresponding Nissl slices

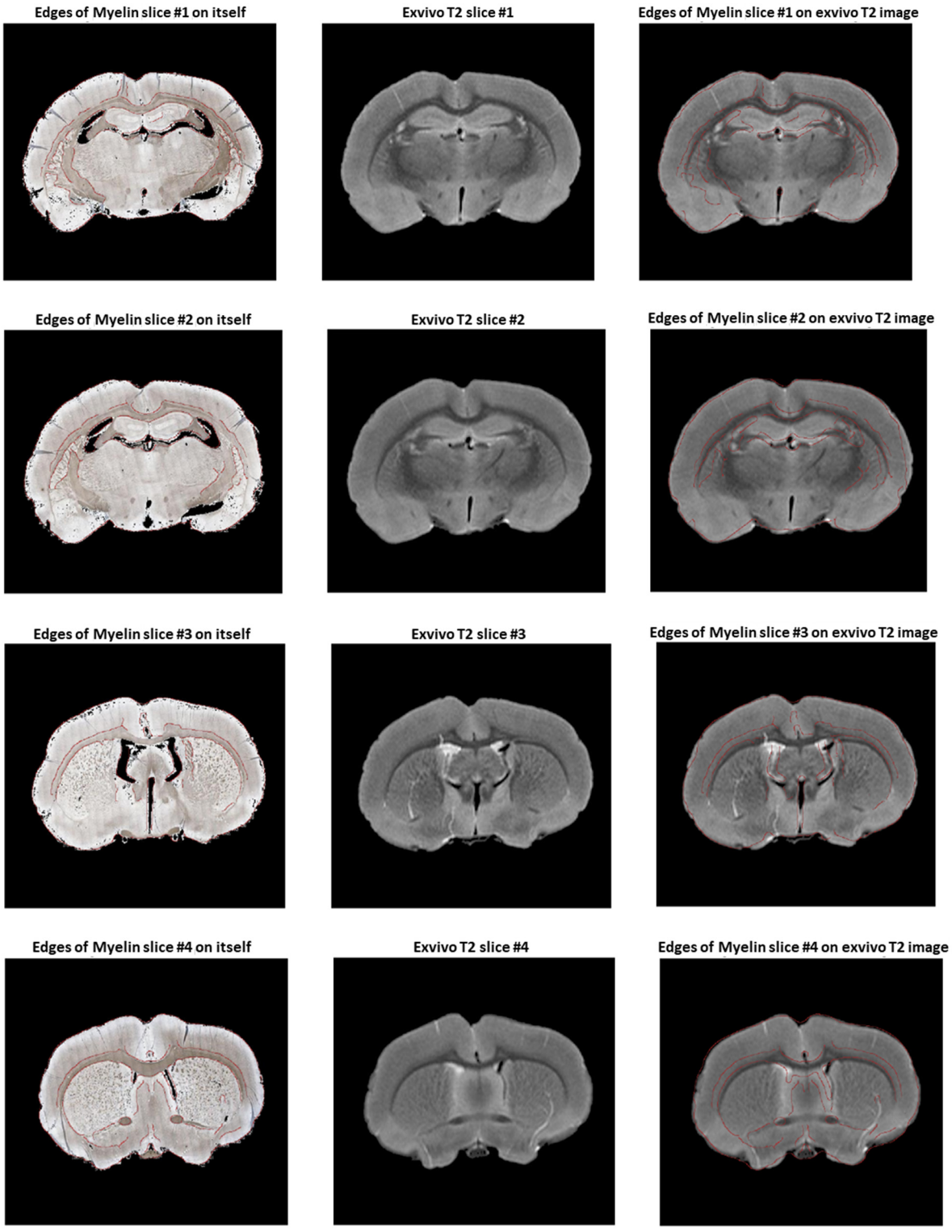

Figure 22: Qualitative evaluation using superimposed edges of the registration between T2 MRI and corresponding Myelin slices