POLITECNICO DI TORINO

Master's Degree in Biomedical Engineering

Master's Thesis

Extending DeepQSM with diffusion MRI

Supervisors:
Prof. Gabriella Balestra, Politecnico di Torino
Prof. Leigh Johnston, Melbourne University

Co-supervisor:
Dr. Steffen Bollmann, The University of Queensland

Candidate:
Giampaolo Tuccio

April 2019

A mia madre e a mia nonna,
per essere sempre al mio fianco
in questo viaggio che é la vita.
A mio padre e a mio fratello,
per essere solido sostegno
ogni giorno, ogni passo.

Abstract

Biological tissue has the capability of locally modifying the magnetic field applied in an MRI scan; this property is called magnetic susceptibility. In order to detect the magnetic field perturbation due to this property, phase images needs to be acquired as in the magnitude images the perturbation causes changes too small to be detected. Mapping magnetic susceptibility distribution is a broad field, recently developed because of possible investigation in iron and myelin changing in clinical and preclinical settings, very important in studying Parkinson's disease, Multiple Sclerosis, along with the neurodegenerative diseases. Unfortunately quantifying the susceptibility property of the brain (QSM - Quantitative Susceptibility Mapping) has drawbacks including dipole inversion, which is an ill-posed problem, and the anisotropic effect which has not been taken into account by the recent gold standard techniques. Recently, machine learning and especially deep convolutional neural networks (DeepQSM), have shown promising results in solving the ill-posed problem, however the anisotropy effect is as yet unsolved. This work attempts to demonstrate how to train a neural network whose knowledge includes information regarding the fibre orientation to solve the ill-posed field-to-source-inversion on *in-vivo* MRI phase data; this was possible using diffusion imaging.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Quantitative susceptibility mapping (QSM) is a recent MRI technique that aims to calculate tissue's magnetic susceptibility distribution within the brain from its perturbing effect on the main magnetic field. This effect is caused by the intrinsic property of the matter to be magnetized if placed into an external magnetic field. The response to this magnetization is a local perturbation that can be aligned with or against the magnetic field. However, the perturbation is so small that until a few years ago, standard low-field MRI could barely detects [10].

The main source of magnetic susceptibility in the brain is iron [40], which is commonly related to brain functions and their change in pathological subjects. Thus its measurement can be used to study several neurodegenerative disease and their development [38], [10], [26], [32], [47], [6], [33].

In order to obtain a quantitative susceptibility map, an image has to be acquired where the signal coming from an MRI sequence needs to be sensitive to the local magnetic field changes. The main challenge in this field is to solve the inverse problem from images that have been pre-processed from raw signal phase acquired with the MRI scan. One can reconstruct the susceptibility source distribution by performing spatial deconvolution of the dipole kernel or by dividing the Fourier transform of the phase image by the Fourier transform of the spherical dipole kernel. Unfortunately, the division in the Fourier domain is an ill-problem because of zeros in the Fourier space. This is a complex problem that has required the development of several techniques such as numerical strategies based [24] or additional measurements based methods like COSMOS [30]. Some drawbacks of these

techniques emerge from:

1. the artifacts introduced when numerical manipulations are implemented in processing the data; each stage of the pipeline is very sensitive to the noise so if it is not well corrected it will be amplified;

2. the requirement of multiple head orientations (see COSMOS in **Section 2.3**) that could be hard to perform in clinical application for human *in-vivo*.

In addition to these, most of methods nowadays used in clinical application does not take into account the consequences coming from anisotropic susceptibility, recorded and highlighted in several papers [43], [47], [46], [25], [29]. This effect arises changing in the frequency shift depending on the orientation of the matter to the magnetic field direction; it is more evident in anisotropic structures (e.g. white matter as it is well known [28], [15]). The result of the anisotropic effect is an incorrect solution of the inverse problem in the voxel where it is expected to be a uniform perturbation response but it is not because of the anisotropy.

Recently the interest in deep learning seemed to be useful and helpful to achieve this problem. It's been demonstrated [38], [48] that U-NET has shown the capability of performing a non-linear mapping from an input space to an output space. The results coming from the literature in this field are very promising and show impressive performance compared to conventional algorithms. In [38] a U-NET model has been trained to solve the inverse-problem. In this work, we investigate whether anisotropy effects can be included in the convolutional neural network that takes into accounts this effect. As far as we know, no previous research has investigated the use of neural networks that involves structural distribution information in order to achieve the susceptibility anisotropy.

The aim of this work is to develop and train a neural network (DeepQSM) to reconstruct isotropic and anisotropic susceptibility effects on the magnetic perturbation, and to include this information in the training phase of the network. Thus, the input of the network should contain information regarding the anisotropy effect, tissue distribution and orientation, along with the phase image.

The following work is divided into:

1. **Theoretical Framework.** It presents a general theoretical overview of the mag-

netic susceptibility in MRI field, its first studies and its relation to different neurological disorders. The inverse problem related to the susceptibility map is discussed and some of the promising techniques used in the last decade are described. Furthermore, the anisotropic effect is described, how it affects the challenging reconstruction and how it might be solved.

2. **Material, Simulations and Methods.** The focus of this thesis yields two main parts: the analysis and generation of a training set containing orientation information for a better reconstruction to the field-to-source problem. The second is the implementation of the U-NET presented by [38], adapting it to a multi-channel input network. Most of the methods required the usage of Python 3.6, to develop most of the coding part and FSL, MRITRIX to pre-process *in-vivo* brain data.

3. **Results.** Here the results of the simulations are shown and presented.

4. **Discussion.** This section provides a discussion of the results of this work and draws attention to the results conducted with *in-vivo* brain data, comparison between the current study and [38], limitation of the study and problems that occurred during the development of this work.

5. **Conclusion.** This last section focuses the attention of further ideas, future solutions and work that might be investigated explaining some of the possible reasons why the results are not the ones expected.

Chapter 2

Theoretical Framework

One area of MRI, that has been rapidly developed in the last decade is that of contrast based on **magnetic susceptibility**, which relates to the tendency of tissue to induce a perturbation when the sample is placed into a magnetic field. This property has long been considered more as a source of artifact and image distortion in magnitude images rather than a benefit. In 1984, Faul et al. [38], [8] started focusing the attention to the effect of tissue susceptibility, but the low field (0.15 T) and low resolution could not enhance an image contrast based on the magnetic susceptibility difference within tissues.

This intrinsic property represents one of the major topics to be investigated in MRI because, as it will be better described in **Section 2.1**, the effect of tissue susceptibility provides information about myelin, white matter composition, iron metabolism and copper accumulation. Therefore it is clear why it is a promising tool for studying **neurodegenerative diseases** such as Huntington's Disease, Multiple Sclerosis, Alzheimer's Disease and Parkinson's Disease or even normal aging. Nowadays, this main interest has been developed thanks to the availability of high field systems (up to 9.4 T for small animal studies)[10], that allow the possibility to interpret even small frequency variations due to heterogeneity of magnetic susceptibility within brain tissue. Across grey and white matter of healthy brain, it varies within a range from about 9.2 to 8.8 ppm. This variation is primarily caused by variations in iron and myelin content, which both have sufficient concentration and susceptibilities sufficiently different from water to contribute substantially. Usually these values are reported in relative terms to that of water($\chi_{H2O} = -9.05 \cdot 10^{-6}$ in SI units) due to its large abundance (70 - 85 % of brain tissue is water) [8], [11].

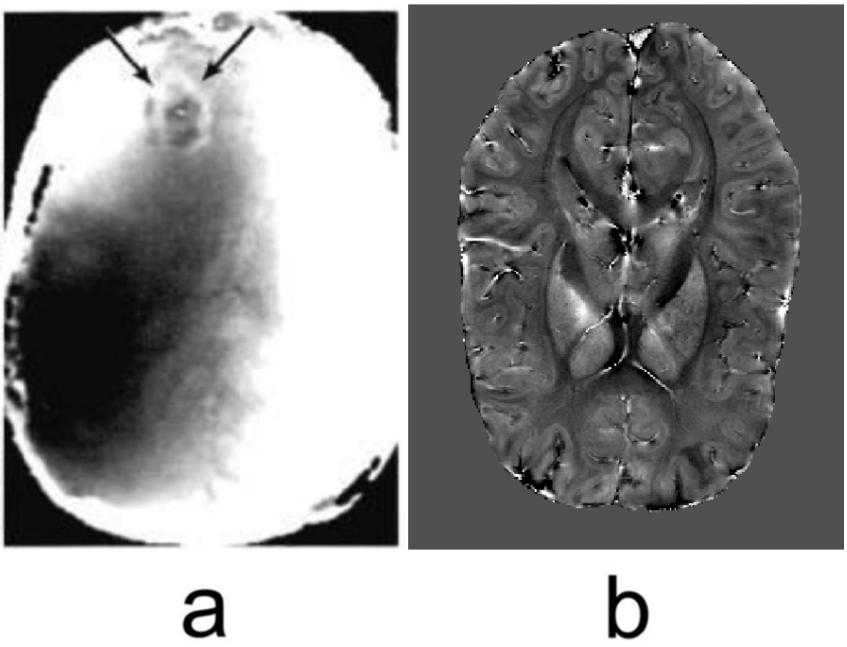


Figure 2.1: Much progress has been made using high field in acquiring phase image. Image (a) shows a brain tumor (arrows) where it is clearly visible the phase shift at its boundary due to the different magnetic susceptibility, however the contrast is very low (image acquired in the 1980's with a field strength of 0.15 T). Image (b) shows the improvement in viewing anatomical details using high field (7 T) scanner at University of Melbourne, after the phase has been unwrapped. Banding and shading effects are due to macroscopic susceptibility heterogeinities like tissue interfesaced Image adapted from [10].

Magnetic susceptibility is defined as the degree of magnetization of a material in response to an applied magnetic field; the material will become magnetized and it induces a pattern of field changes (the so-called *demagnetization* field) that depends on the chemical structure of the sample, the shape and orientation respect to the magnetic field, and generally it extends beyond the object's border. The effects of tissue susceptibility variations on the magnetic field can be studied with susceptibility weighted MRI through acquisition of multiple gradient echo signals (see **Section 2.2**, from which both amplitude and frequency are extracted). Tissue susceptibility can affect both images according to the following equation:

$$f_L(\vec{r}) = \frac{\gamma}{2\pi}[B_0 + \Delta B_z(\vec{r})] \tag{2.1}$$

where f_L is the Larmor frequency, γ is the gyromagnetic ratio (42,57 MHz/T for water ^{1}H proton, B_0 the amplitude of the static field of the MRI scanner and ΔB_z the field shift created by the magnetization M of the object exploited in 2.2. Mathematically, magnetic susceptibility is a dimensionless constant of proportionality as follows:

$$\vec{M} = \chi \cdot \vec{H} \tag{2.2}$$

where $\vec{M}$ represents the induced magnetization, $\vec{H}$ is the applied magnetic field intensity in $A \cdot m^{-1}$ and χ is the scalar quantity for the susceptibility. In free space, where there is no magnetization, the magnetic field field can be expressed equally well by $\vec{B}$ or $\vec{H}$ which are linearly related by $\vec{B} = \mu_0 \cdot \vec{H}$. In magnetized matter, the linear relationship is not valid anymore and the relation stands in the following equation:

$$\vec{B} = \mu_0 \cdot (\vec{H} + \vec{M}) = \mu_0 \cdot (1 + \chi) \cdot \vec{H} \tag{2.3}$$

If the magnetic field in the material is strengthened by the induced magnetization, the material is defined as **paramagnetic** and the χ value is greater than 0, otherwise it is called **diamagnetic**. Brain tissues susceptibility values have been measured with NMR and they are usually referenced to water because of the high quantity in biological tissues. In the upcoming text χ values will be always referenced to water, as *relative* susceptibility values.

2.1 Biophysical mechanism of phase contrast

The most relevant biophysical sources of different susceptibility values *in-vivo* are myelin, water, iron and calcium [40], [10], [47]. Calcium has a lower, so more diamagnetic, susceptibility than water, favouring differentiation between calcium and blood or blood products in brain lesions. Regarding iron, it is present in two configurations, as storage iron (non-heme iron), such as ferritin and hemosiderin, or it is bound to heme proteins, such as in haemoglobin (heme iron) as explained in [47]. So far most of the techniques used to map iron deposition in the brain are invasive, while susceptibility measurements mapped by MRI is a promising non-invasive technique to study iron in various neurological and

psychiatric disorders.

Myelin represents a significant contributor to tissue magnetic susceptibility in the Central Nervous System (CNS); it is predominantly present in white matter being a lipid-rich substance that forms a multi-layered membrane structure composed of lipid and proteins around the axon, in order to improve action potential propagation [27]. The contribution of the myelin to the measured bulk susceptibility has been proved in a number of studies [26]. In 2011, Liu et al. [36], [26] showed that the phase and susceptibility contrast between grey and white matter were reduced by more than 90% in shiverer mice (genetically modified in order to present a hypo myelinated condition in the CNS) compared to control mice. In both situations, other MRI tools like fractional anisotropy (FA), coming from Diffusion Tensor Imaging (DTI), was altered only slightly due to the same axonal structures, while susceptibility weighted imaging could clearly map the difference.

However, a tissue's chemical and molecular composition are not the only reasons that affect magnetic susceptibility values; the microscale arrangement may also influence the measured phase shift within a voxel; microstructural effects can be assigned to two contributions: (i) tissue microstructure at the cellular and subcellular level, not directly detectable by MRI, and (ii) molecular susceptibility anisotropy [26]. Both parts play an important role in susceptibility measurements in highly aligned structures such as WM fibre bundles, renal tubules or muscle fibres. A schematic illustration of WM fiber is represented in figure 2.2.

2.1.1 Potential of magnetic susceptibility contrast to solve clinical issues

Some MRI techniques, such as SWI, phase imaging and R_2^* mapping are currently used in the clinical routine used to assess selected pathological susceptibility variations, as abnormal venous vessels, microlesions or abnormal iron content [18], [12]. Despite QSM, these techniques are mostly qualitative and not quantitative, since they measure the susceptibility indirectly and suffer from reduced CNR (Contrast to Noise Ratio) of deep GM and low sensitivity. QSM represents instead a superior tool because it overcomes several of these limitations and enables the differentiation of iron and calcium deposits

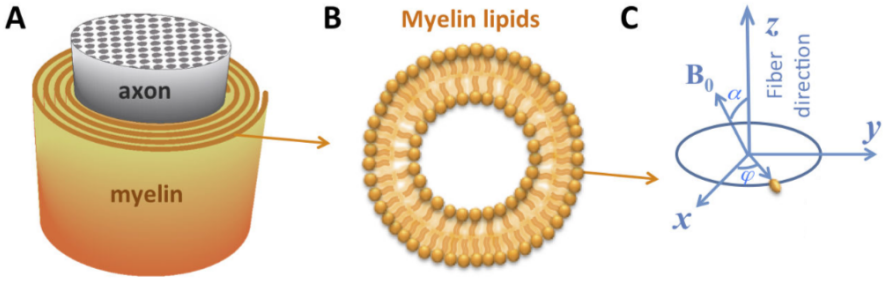


Figure 2.2: Schematic representation of the axon structure highlighting the myelin sheath around an axon (A). (B) Radial alignment of myelin lipid molecules. (C) The orientation between the applied magnetic field and the fiber direction, which plays an important role in anisotropic susceptibility. Figure adapted from [26].

[8]. Generally, the main clinical applications of QSM might be:

1. Anatomical imaging of the human brain: thanks to its high sensitivity to myelin and iron in combination with high spatial resolution given by high field magnetic strength, it would be possible to have a detailed view of the brain morphology. Great potential might be carried by QSM in deep brain surgery where high and spatially well-localized contrast is essential. The only limitation of QSM could be represented by orientation-dependent effect caused by high aligned structures like WM fibers in corpus callosum. (More about this in **Section 2.3.2**).

2. Traumatic brain injury (TBI): this is an injury to the brain caused by an external force. Common causes include falls, car accidents, assault or being struck by objects such as might occur during sport. It is often related to (micro)hemorrhages resulting from blood-brain barrier permeability changes or injuries of small vessels. Several studies have associated the presence and location of TBI-related hemorrhages to specific neuropsychological deficits [16], [3]. Previous studies have shown greater accuracy from susceptibility-techniques than CT and conventional MRI tools.

3. Brain tumors: several studies [9] have now proved the correlation between calcification and brain tumor onset (recurrent glioblastoma [1] is just one of this). Imaging of

[1]Glioblastoma is a common high grade brain tumor, considered the most aggressive cancer that begins

variations in magnetic susceptibility revealed advantages in depicting the internal microarchitecture of the venous vasculature.

4. Neurodegenerative diseases [8]: some of the most studied diseases have shown strong correlation iron level variations *in-vivo*. Several studies reported iron elevation in specific brain regions [14], [13], [23] related to Parkinson's disease - high iron concentration in the substantia nigra (SN) - Alzhaimer's disease - with an increased magnetic susceptibility in deep GM nuclei, particularly putamen - and multiple sclerosis, where the demyelination is the main feature of the disease.

2.2 Signal Acquisition

Since the phase shift dependence of the magnetic susceptibility, any MR sequence that can detect phase variation during the data acquisition can be used in order to acquire a proper image. Particularly used is the gradient-recalled echo (GRE) MRI because phase contrast images with high SNR can be obtained from this technique since it is very sensitive to resonance frequency variations (and so to susceptibility variations). The GRE signal phase $\Phi(\vec{r}, TE)$ is equal to:

$$\Phi(\vec{r}, TE) = \Phi_0(\vec{r}) + 2\pi\Delta f(\vec{r}) \cdot TE = \Phi_0(\vec{r}) + \gamma\Delta B_z(\vec{r}) \cdot TE \tag{2.4}$$

where TE is the echo time, the time between the first excitation and the signal reading during the rephasing gradient. It has been shown that optimum phase contrast is achieved for a TE equal to the tissue's effective transverse relaxation time T_2^*.

QSM requires some essential pre-processing steps, after image acquisition, in order to solve the ill-posed problem starting from accurate phase data. The target of this text is not to focus on these techniques, but the following are the main steps:

1. **Coil Combination**: Since there is more than one coil, or channel, is used in ultra-high field MRI, when one must combine the phase from each coil. There is an offset in the measurement of each coil;

within the brain

2. **Phase unwrapping**: The measured phase signal can only take on values in a 2π range, but the original phase signal can take on any possible value! This leads to so-called phase wraps, that require an unwrapping algorithm in order to estimate the true phase;

3. **Background field removal**: Some external sources contribute to the demagnetization field so that the measurement is a sum of the magnetic field change due to internal sources and magnetic field change caused by the background, that needs to be removed;

4. **Field-to-susceptibility inversion** in order to achieve and reconstruct the original source of perturbation;

2.3 What is QSM?

The post-processing technique adopted to map and detect magnetic susceptibility distribution from MRI phase measurements is called Quantitative Susceptibility Mapping (QSM). It is based on the assumption that the field pattern can be found by (spatially) convolving $\vec{M}$ with the magnetic field generated by a point dipole (an infinitely small magnet with unit magnetization). This is the so-called forward problem. For a simple case from (sub)atomic scale dipole, an infinitely extending uniform magnetization M this dipole convolution results in an uniform field shift $\Delta B = \mu_0 M$ (μ_0 is the magnetic permeability of free space). An additional field shift can be taken into account if one considers an outer boundary (e.g. sphere, cylinder) geometrically simple. This results a "demagnetization" or "shape" factor which for a sphere is represented by a subtraction of a factor $\frac{1}{3}\mu_0 M$. Unfortunately, χ and M are defined macroscopically but actually they result from (sub)atomic scale dipoles, so further complications must be taken into account. Commonly used is the "Lorentz sphere" approach, introduced in 1909 by Hendrik Lorentz [31], based for simplicity on a spherical Lorentz surface. It can be defined arbitrarily, because it is only a mathematical concept to facilitate the magnetic field calculation and it has no physical foundation, but it cannot be arbitrarily small. The surface must be chosen big enough in order that the total field at location $\vec{r}$ be modelled as resulting from a continuous medium with magnetization $\vec{M}$. Lorentz suggests the surface be larger

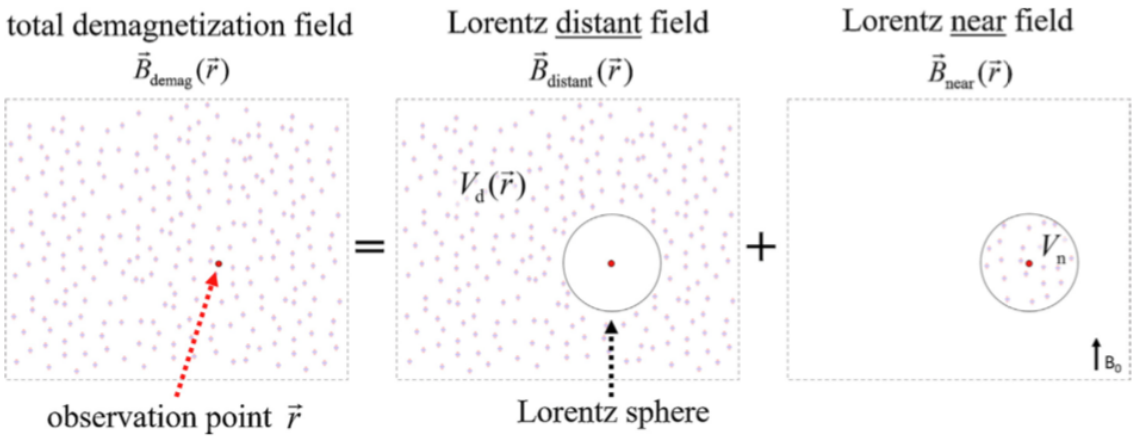


Figure 2.3: Illustration of the Lorentz model. On the left, the total demagnetization field $\vec{B}_{demag}(\vec{r})$ that can be decomposed into a distant field $\vec{B}_{distant}(\vec{r})$ (middle) and a near field $\vec{B}_{near}(\vec{r})$ (right) resulting from a virtual Lorentz cavity around the observation point $\vec{r}$. Figure adapted from [41].

then the mutual distance of nearby magnetic moments. Typically, in MRI literature it is assumed that magnetic moments are randomly distributed in the sample so that the field contributions within the cavity average to zero. This does not happen if the arrangement of dipoles is more regular like in white matter [45], [42]. Figure 2.3 shows an illustration of the Lorentz approach.

2.3.1 Dipole Kernel

The heterogeneous magnetic susceptibility values in the brain can lead to local magnetic field differences and thus resonance frequency variations over the tissue. Such shifts have two major effect; first, it regards the opportunity to map anatomical details inside the brain thanks to the magnetic properties of the brain tissue (detected using gradient-recalled echo (GRE) phase images, which are sensitive to shifts resonance frequency); second, a shift resonance frequency may lead to undesired effects such as image distortion. The goal of susceptibility imaging is to exploit the former while minimizing the effects of the latter.

QSM can be achieved in three main steps: (i) estimating the magnetic field distribution from raw MRI phase data, then (ii) removing the *background field* contribution given by the source outside the field-of-view, such as skull, paranasal sinus, human torso,

which are also exposed to the static magnetic field. Finally, (iii) the ill-posed problem has to be solved from field perturbation to magnetic susceptibility. But first, it is important to describe the forward problem in order to understand what are the unknowns and why certain techniques have been chosen to solve the inverse problem: It can be demonstrated that:

$$\Delta B_z = \mu_0 \int M_z(\vec{r}') \cdot d_z(\vec{r} - \vec{r}') d^3\vec{r}' \tag{2.5}$$

introducing the Green's function[2], a dipole kernel based on Lorentz sphere:

$$d_z(\vec{r}) = \frac{1}{4\pi} \cdot \frac{3\cos^2(\theta_{3D}) - 1}{|\vec{r}^3|} = FT^{-1}\left\{\frac{1}{3} - \frac{k_z^2}{k_x^2 + k_y^2 + k_z^2}\right\} \tag{2.6}$$

where μ_0 is the magnetic permeability of free space, θ_{3D} is the azimuth angle in the spherical coordinate system between z direction (where the main field is oriented in an MRI system) and the dipole, $M_z(\vec{r}')$ is the magnetization, $\vec{r}$ and $\vec{k}$ are the coordinate positions in the image and k-space, respectively, and in particular $r^2 \equiv x^2 + y^2 + z^2$.

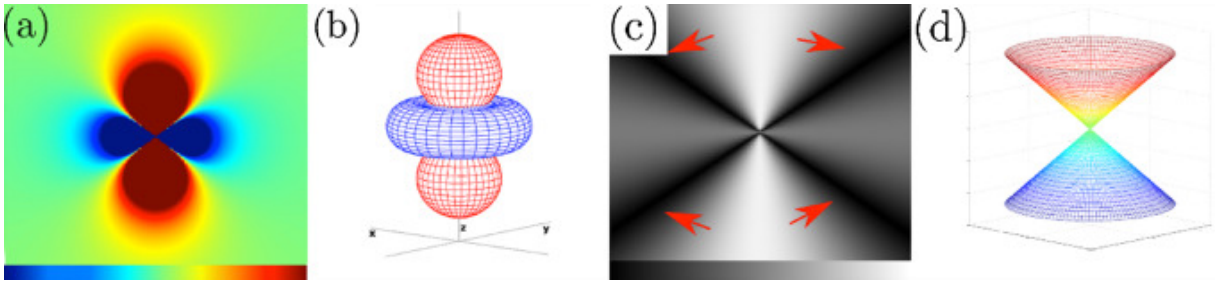


Figure 2.4: Unit dipole response (Green's function) in the image space, (a,b) and in the k-space (c,d), respectively. (a) and (c) are the sagittal sections, while (d) shows the zero cones in k-space spherical model kernel. Adapted from [41].

Assuming that $|\chi| \ll 1$, as in brain tissue, it follows that $M_z(\vec{r}') \approx \chi(\vec{r}') \cdot \mu_0^{-1} \cdot B_0$ in equation [2], in order to provide the link between the magnetic field variations and the

[2]Green's function has been used to describe the dipole of a point source. Based on the Lorentz Sphere assumption (introduced by Lorentz in 1909), this concept explained that the local electrical field in an environment full of electrically polarized particles can be seen as the sum of two different contributions, the distant field and the near field.

susceptibility distribution, which can also be written in the following form:

$$\Delta B_z(\vec{r}) = B_0 \int \chi(\vec{r'}) \cdot d_z(\vec{r} - \vec{r'}) d^3\vec{r'} = FT^{-1} \left\{ \chi(\vec{k}) \cdot d_z(\vec{k}) \right\} \cdot B_0 \tag{2.7}$$

Moreover, $\Delta B_z(\vec{r})$ has to be determined from MR measurements. The quantity $\frac{\Delta B_z(\vec{r})}{B_0}$ is also referred to as *Relative Difference Field* (RDF) [41]. It's well known that the local magnetic field of each nucleus induces the Larmor frequency of that nucleus as in **Equation 2.1**. Hence, the magnetic field variations induce frequency shifts compared to the scanner's frequency $f_R \approx \frac{\gamma}{2\pi} B_0$ as it follows:

$$\Delta f(\vec{r}) = f_L(\vec{r}) - f_R = \frac{\gamma}{2\pi} \Delta B_z(\vec{r}) \tag{2.8}$$

The calculation of $\chi(\vec{k})$ from equation (4) represents the **inverse problem** of the QSM; this complex expression becomes simple, computationally speaking, when calculated in the Fourier domain as a fraction between the Fourier transform of the magnetic field variation distribution (obtained by (6)) and the Green's function in k-space, then inverting the result one has achieved the problem in the image-space. The problem comes when dealing with the kernel in the Fourier domain because of the 0 values at the denominator; thus, in mathematical terms, this means that the inverse problem is an *ill-posed* problem with no unique solutions and it is impossible to retrieve the corresponding spatial frequency components of χ values from measurements of the RDF.

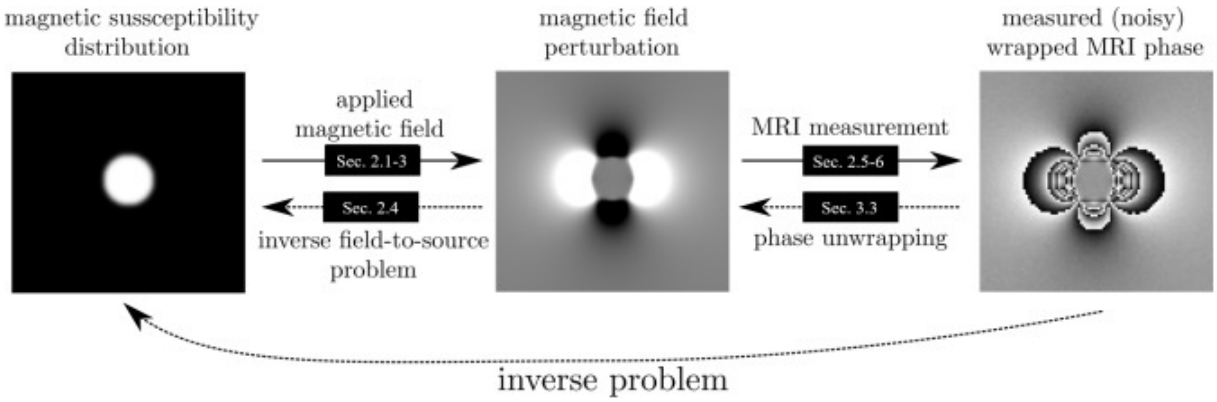


Figure 2.5: Schematic illustration of the inverse problem. Adapted from [41].

Many techniques have been proposed in order to solve the undetermined problem,

such as iterative solution using regularization techniques with implicit or explicit prior information. One of these is **TKD**, Truncated K-space Division. According to this technique it suggests thresholding the dipole kernel in k-space to remove 0 values that cause the *ill-posed* problem. The threshold should be decided depending on SNR: a smaller threshold is good enough for a good image with high SNR. The main disadvantage of TKD is the noise introduced after division by RDF by the thresholded dipole.

The most successful technique is **COSMOS** [30]. Instead of dealing with regularization strategies, COSMOS (Calculation Of Susceptibility through Multiple Orientation Sampling) tried to achieve the solution through multiple orientation measurements and finally solving a linear equation system that comprised all the data coming from multiple orientations (over-determining the problem). The drawbacks of this technique are the long acquisition time and the difficulties related to the clinical applicability of different required head orientations. Nevertheless, COSMOS is been considered as a gold standard for several years.

2.3.2 Anisotropic susceptibility

In the past few years, this view has radically changed, as extensive research has demonstrated that the magnetic susceptibility of brain white matter depends on the direction of the main magnetic field relative to the axons of neurons [15]. COSMOS assumes that in each n^{th} orientation the magnetic susceptibility values is the same within the same brain voxel (*isotropic* assumption). But, it has been possible to see from test observations [41] the orientation-dependence of magnetic susceptibility in the mouse central nervous system. He and Yablonskiy [15] were the first to demonstrate the effect of the orientation on the frequency perturbation.

It is well known the frequency shift due to magnetic susceptibility effects assuming the Lorentz sphere model, and this can be described as the sum of 2 contributions:

$$\frac{\Delta f}{f_0} = A \cdot \chi + \frac{4}{3}\pi \cdot \chi \tag{2.9}$$

where the first term describes effects on the object general external shape and and the second term shows the effect of the Lorentzian sphere.

He et al. [46] demonstrated the difference result one should expect when anisotropy effect is taken into account. Indeed, assuming a longitudinal structure such as a WM axon fiber parallel to B_0, if one would use a standard Lorentzian sphere approximation, one would expect that the magnetic susceptibility induced frequency shift would be different from zero from 2.9 in the surrounding space of the fiber, but this result is not correct. From [17] the magnetic field effect in this case should be neglectable due to the longitudinal orientation of the above-mentioned structure with respect to the B_0.

According to [46], the model can be addressed evaluating the frequency shift as following:

$$\left.\frac{\Delta f}{f_0}\right|_{sphere} = 2 \cdot \pi \cdot \chi_a \sin^2(\theta), \tag{2.10}$$

where θ is the angle between the axonal direction and the external field B_0. Of course, this frequency shift is additional to the frequency shift due to the isotropic part that can still be described by the second term of 2.9:

$$\Delta f = \Delta f|_{isotropic} + \Delta f|_{anisotropic} \tag{2.11}$$

This result is called the *Generalized Lorentzian Approach* [15].

Lorentz cylinder

Subsequently an alternative model has been proposed, [46], [47], [43], where a Lorentz cylinder addresses the orientation-dependence issue and calculates correctly the magnetic field changing when the anisotropic behaviour occurs. Thus, the frequency shift depends on the spatial arrangement (architecture) of the tissue in the brain at the cellular and sub-cellular level and it can be better described by an infinitely long Lorentz cylinder. Yablonskiy and He [46] point out that the frequency shift (f) can be described as follows:

$$\left.\frac{\Delta f}{f_0}\right|_{cylinder} = \frac{1}{2}\chi_{long} \sin^2(\alpha) \tag{2.12}$$

where χ_{long} is the contribution to the tissue volume magnetic susceptibility from

orientated longitudinal structures, α is the angle between the direction of the longitudinal structure and B_0. This model takes into account both contribution, isotropic and anisotropic, expressed by a Lorentz cylinder depending on the structure that appear in each voxel. Of course, the frequency shift depends on both terms if one would include the cylinder model, and the resulting equation would be:

$$\left|\frac{\Delta f}{f_0}\right| = \frac{1}{3}\chi_{iso} + \frac{1}{2}\chi_{long}\sin^2(\alpha) \tag{2.13}$$

[35] investigated a similar approach that employs the spherical and cylindrical models of susceptibility effect. For each voxel it has been measured how uni-directional is the diffusion according to FA (Fractional Anisotropy), where high FA corresponds to diffusion occurring along a single direction (**anisotropy**) and low FA corresponds to diffusion occurring equally in all directions (**isotropy**). From a chosen threshold, the voxels are classified as Lorentz sphere or Lorentz cylinder.

2.3.3 Susceptibility Tensor Imaging

The anisotropy of magnetic susceptibility led to the proposal of Susceptibility Tensor Imaging (STI), that provides a mathematical model of the anisotropy of magnetic susceptibility in each voxel [13]. The idea is very similar to the tensor model used in Diffusion Tensor Imaging (DTI), in which the tensor describes the pattern of water diffusion along 3 main axes within each voxel. The tensor is a 3x3 symmetric matrix describing an ellipsoid with three unit eigenvectors (ϵ_1, ϵ_2 and ϵ_3) and corresponding eigenvalues values (λ_1, λ_2 and λ_3).

If the tensor's three eigenvalues are equal [43], the voxel can be assumed to be isotropic (a sphere), otherwise one would have a rank 2 (or second-order) 3 x 3 tensor symmetric matrix with 6 unknown components to be calculated.

$$\begin{bmatrix} \chi_{11} & \chi_{12} & \chi_{13} \\ \chi_{12} & \chi_{22} & \chi_{23} \\ \chi_{13} & \chi_{23} & \chi_{33} \end{bmatrix} \tag{2.14}$$

If eigenvalue decomposition is performed [43], three principal susceptibilities can

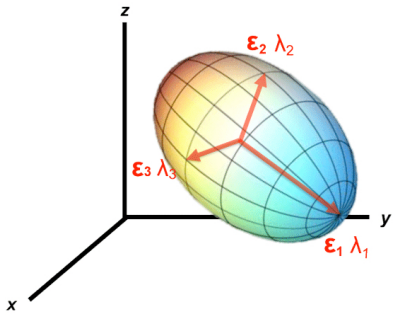


Figure 2.6: The susceptibility/diffusion ellipsoid represented with its eigenvectors $(\epsilon_1, \epsilon_2, \epsilon_3)$, with corresponding length $(\lambda_1, \lambda_2, \lambda_3)$, the eigenvalues [2].

be defined (χ_1, χ_2, χ_3) each with a correspondent eigenvector, of which the biggest, the principal eigenvector points toward the direction that exhibits the largest magnetic susceptibility (paramagnetic) and the minor eigenvector points in the direction with the most negative (diamagnetic) susceptibility. Including isotropic and anisotropic contribution, the susceptibility tensor related to a voxel would be like the following expression:

$$\begin{bmatrix} \chi_{||} & 0 & 0 \\ 0 & \chi_{\perp} & 0 \\ 0 & 0 & \chi_{\perp} \end{bmatrix} = \begin{bmatrix} \chi_I & 0 & 0 \\ 0 & \chi_I & 0 \\ 0 & 0 & \chi_I \end{bmatrix} + \begin{bmatrix} \chi_A & 0 & 0 \\ 0 & -\chi_A/2 & 0 \\ 0 & 0 & -\chi_A/2 \end{bmatrix} \tag{2.15}$$

Where χ_I and χ_A are scalar values describing the isotropic and anisotropic components of a cylindrical symmetric susceptibility tensor,$\chi_{\perp}$ and $\chi_{||}$ are the magnetic susceptibility of the sample perpendicular and parallel to the principal axis.

In QSM, the main parameters derived from the susceptibility tensor are the mean magnetic susceptibility (**MMS**):

$$MMS = \frac{\chi_1 + \chi_2 + \chi_3}{3} \tag{2.16}$$

and magnetic susceptibility anisotropy (**MSA**), that reflects the same meaning of **FA** (Fractional Anisotropy) in DTI:

$$MSA = \chi_1 - \frac{\chi_2 + \chi_3}{2} \qquad with \quad \chi_3 \leq \chi_2 \leq \chi_1 \tag{2.17}$$

Two possible approaches have been proposed in the literature, a k-space-based and an image-space based method [26], [28]:

1. **k-space approach**: the susceptibility tensor is calculated in the frequency domain as:

$$\delta(k) = FT(\frac{\Phi}{2\pi TE\gamma\mu_0 H_0}) = \left[\frac{1}{3}\hat{H}^T FT(\hat{\chi})\hat{H} - k^T\hat{H}\frac{k^T FT(\hat{\chi})\hat{H}}{k^2}\right] \tag{2.18}$$

where Φ is the MR measurement, γ is the gyromagnetic ratio, μ_0 is the vacuum permeability, H_0 is the magnitude of the applied magnetic field, $\hat{H}$ is the unit vector of the applied field, k is the spatial frequency vector, and it can be expressed as a weighted sum of six distinguishable susceptibility tensor components in k-space:

$$\delta(k) = a_{11}\chi_{11}(k)+a_{12}\chi_{12}(k)+a_{13}\chi_{13}(k)+a_{22}\chi_{22}(k)+a_{23}\chi_{23}(k)+a_{33}\chi_{33}(k) = \mathbf{A}\chi(k) \tag{2.19}$$

Thus, the tensor map can then be estimated by solving a system of linear equations using least squares methods, requiring measurement of at least 6 different head orientations.

2. **the image-space approach**: In this case, the tensor matrix is estimated by solving the optimization:

$$\min_x(\|O_1\|_2^2 + \|O_2\|_2^2 + \cdots \|O_n\|_2^2) \tag{2.20}$$

where $O_i = A\chi(r) - b_i$ with $i = 1, \cdots, N$ indicating the i^{th} head orientation. This approach is computationally more intensive but has the advantage of incorporating constraints based on image space features, for example, assumptions of isotropic behaviour in voxel represented by CSF or white matter; it has been demonstrated that this regularization improve the image quality. Indeed, one cannot assume any constraints in image space in the k-space approach.

However, considering the anisotropy assumption, it has been demonstrated [26] that the bulk susceptibility anisotropy in white matter, modelling as cylindrical shells with lipid

molecules radially aligned within these shells, can be assumed as follows:

$$\chi = \mathbf{M} \cdot \hat{\mathbf{H}}/H_0 = f_{lipid} \left(\frac{x_{//} - x_{\perp}}{2} \right) sin^2\alpha + \chi_0 \tag{2.21}$$

where $\hat{\mathbf{H}}$ is the unit vector of the applied field, H_0 is the applied field strength, χ_0 is the baseline isotropic susceptibility difference.

2.4 Diffusion weighted image (DWI)

The previous study [25] provided an important opportunity to advance the understanding of a relationship between Diffusion images and Susceptibility images. Although several studies [45], [25], [29] have defined an evident similarity between the two techniques - [25] attempts to show how similar are the DTI-based-tractography and STI-based, although the gold-standard techniques remains DTI because of its high resolution and detailed detection of the fibers.

As a consequence of this, it is important in the aim of this work describing what is DWI. In this case, image contrast is determined by the random microscopic motion of water protons, which is easily detectable since its abundance in the human body (60%-70%) [4]. The physical process behind DWI is the diffusion which was first described by Robert Brown who observed pollen moving randomly under his microscope in 1827. Diffusion is then the random Brownian motion of the water molecules driven by thermal energy produced by body temperature, in figure 2.7. In a perfectly homogeneous medium, diffusion is random and isotropic, and leads to equal probability in all directions. But in a complex environment such as the human body, the movement is strictly correlated with the tissue boundaries in anisotropic structures. For this reason it has a great potential in imaging and can provide good knowledge into cell physiology.

2.4.1 DWI image acquisition: Pulsed Gradient Spin Echo

To sensitize MRI to diffusion, the homogeneity is varied linearly by a pulsed field gradient. Since precession is proportional to the magnetic strength, the protons begin to precess at different rates, resulting in dispersion of the phase and signal loss, which is spatially

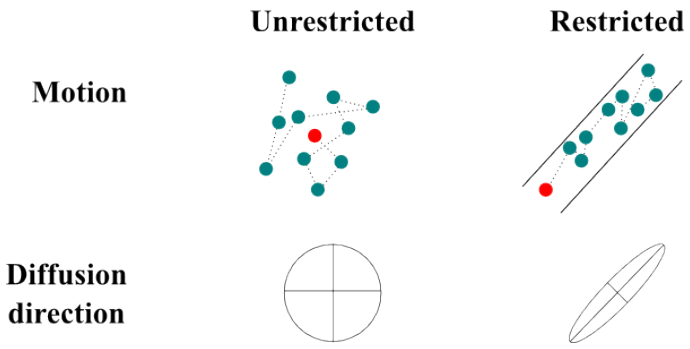


Figure 2.7: The movement of water can be described in general by Brownian motion, or simply speaking a random walk. This happens due to collisions with other water molecules over time and it depends, as shown in figure, on the tissue's structure and the degree of freedom of the molecules. Adapted from [1].

encoded. Another gradient pulse is applied in the same magnitude but with opposite direction to refocus or rephase the spins. The refocusing will not be perfect for protons that have moved during the time interval between the pulses, and the signal measured by the MRI machine is reduced. Figure 2.8 illustrates how moving molecules can affect the phase shift after the rephasing.

The reduction in signal due to the application of the pulse gradient was first described by Stejskal and Tanner [22] through the following equation:

$$S = M_0 e^{-TE/T_2} exp\left[-\gamma^2 G^2 \lambda^2 \left(\Delta - \frac{\delta}{3}\right) D\right] \tag{2.22}$$

where the term that multiplies the exponential is the non-diffusion weighted signal, M_0 is the initial signal strength at t = 0, TE is the echo time, T_2 the transverse relaxation time, λ is the gyromagnetic ratio, G is the strength of the gradient pulse, δ is the duration of the pulse, Δ is the time between the two pulses and D is the apparent diffusion coefficient. D is what is modified if there are boundaries restriction that stops freely moving of water.

However, since the diffusion is only seen in the same direction of the gradient, multiple acquisition has to be performed in order to get a reasonable good image; thus,

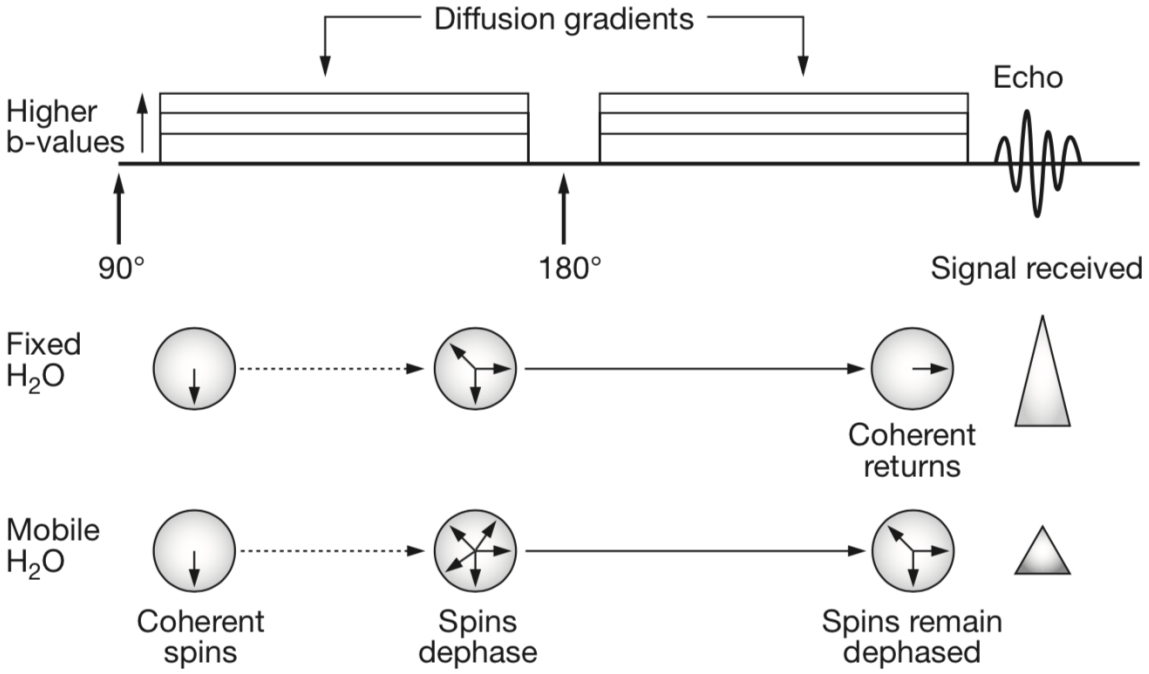


Figure 2.8: Schematic illustration of the diffusion-weighted imaging sequence and the cause of signal loss by mobile protons. The image is produced by the application of two gradients on either side of the refocusing 180 pulse. Increasing gradient heights increases the b-values of images produced. Water molecules that have moved during the application of the first gradient will experience greater phase shifts, resulting in a net loss of signal. In molecules that have no net movement, there is no phase shifts after the rephasing, so the signal is not lost. Figure adapted from [37]

the equation 2.22 becomes

$$S_j = M_0 e^{-TE/T_2} e^{-b g_j^T D g_j} \tag{2.23}$$

where j refers to the direction, b is usually referred to as b-value and indicates the diffusion characterization, selected by the operator prior to imaging. Typical values are from 0 to 1000 s/mm^2, b_j is the gradient direction that characterizes a certain direction j, and D in this case is the Diffusion Tensor, a symmetric 3x3 matrix where diagonal elements are proportional to diffusion displacement variations along the 3 main directions (x, y, and z). Off-diagonal elements depends on the correlation of displacement along the

3 directions. In total it contains 6 unknowns. Once D has been estimated, it is possible to know the direction of diffusivity in each voxel computing first the eigenvalues and eigenvectors. If the eigenvalues are similar, the related voxel is considered as isotropic, otherwise, if one eigenvalue is greater than the other two, the related eigenvector (known as **principal eigenvector**) describes the dominant direction. Useful parameters extracted from the eigenvectors are mean diffusivity and fractional anisotropy (FA), in DWI as well as SWI.

2.5 Deep Learning and Convolutional Neural Network (CNN)

Deep learning refers to a subfield of machine learning that is based on learning levels of representations, corresponding to a hierarchy of features, factors or concepts, where higher-level concepts are defined from lower-level ones. Deep learning is learning multiple levels of representation and abstraction, helps to understand the data such as images, audio and text. The concept of Deep Learning comes from the study of the multilayer perceptron[3] which contains more hidden layers is a Deep Learning structure. In the late 1980s, the invention of Back Propagation algorithm used in Artificial Neural Network was a promising innovation for machine learning field that since that date it has grown up and has been applied to several fields. A class of deep learning neural network is the Convolutional Neural Network (CNN or ConvNet) which represents an efficient recognition algorithm widely used in pattern recognition and image processing. It has many features such as simple structure, less training parameters and adaptability. CNNs were inspired

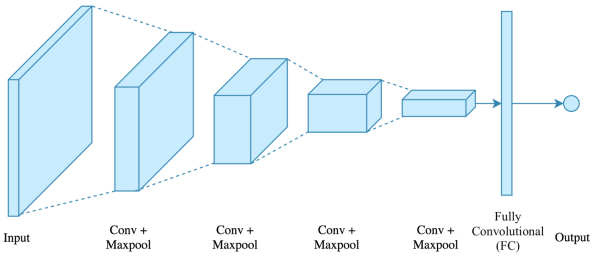


Figure 2.9: Example of CNN for 2D image classification.

[3]In machine learning, the perceptron is an algorithm for supervised learning of binary classifiers. It is a single layer neural network and a multi-layer perceptron is called Artificial Neural Network.

by biological processes [34], [33] in which the connectivity pattern between neurons resembles the organization of the visual cortex. Individual cortical neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. The receptive fields of different neurons partially overlap and cover the entire visual field.

In this work it will be exploited the training and the usage of a U-NET (or Fully Convolutional Neural Network), a convolutional neural network that was developed for biomedical image segmentation. Despite a conventional CNN, the main innovation of U-NET is to supplement a usual contracting network by successive upsampling layers. Hence these layers increase the resolution of the output. In figure 2.10 a schematic and general representation of a U-NET. The U-NET model can be described as composed by two different parts: (1) the first part, denoted as feature extraction layer, is the encoder where the convolution is performed and followed by a maxpool downsampling. Here the input of each neuron is connected to the local receptive fields of the previous layer, and extracts the local feature. The purpose is to to encode the input image into feature representations at multiple different levels. (2) In the second part, denoted as UP, is the decoder of the model and it consists of upsamples and concatenations followed by regular convolution operations. It is based on the idea of combining matching levels: since upsampling is a sparse operation it needs a good prior from earlier stages to better represent the localization. After each upsampling, each block is concatenated to the previous one coming from the left side of the model.

It has to be denoted that every feature map in each level has the same dimension of the input of that later and the weight of the neurons in the feature map are equal.

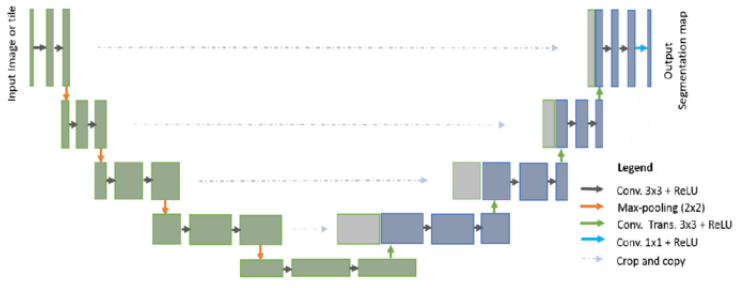


Figure 2.10: Generic model of U-NET for a 2D image segmentation. It is possible to see the downsampling (contracting or convolution) path and upsampling (expanding or deconvolution) path.

Besides, since the neurons in the same mapping feature share weight, the number of free parameters of the network is reduced.

2.5.1 The convolution step

ConvNets derive their name from the convolution operator. The primary purpose of Convolution in case of a ConvNet is to extract features from the input image. Convolution preserves the spatial relationship between pixels by learning image features using small squares of input data. The convolution is usually performed sliding a 3D matrix called "filter", "kernel" or "feature detector" over the image input and computing the dot product. The result is the so-called the "Convolved Feature" or "Activation Map" or the "Feature Map". Though, different filters will produce different feature maps from the same image, see Figure 2.11 for some example kernels.

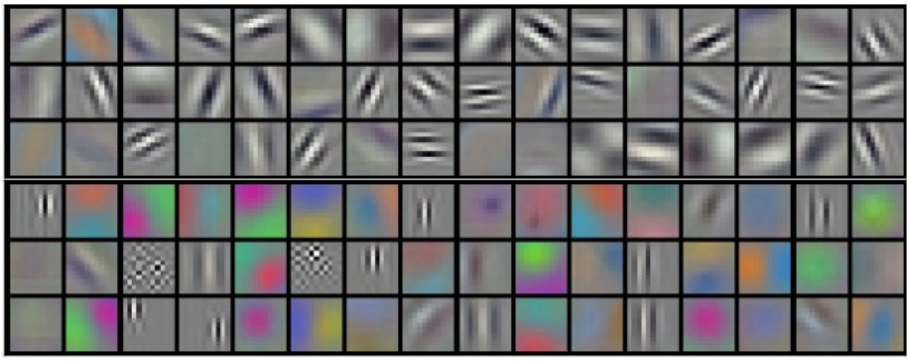


Figure 2.11: Example filters used in the convolution learned by [20]. Image adapted from [21]

2.5.2 The Pooling step

Spatial Pooling (also called subsampling or downsampling) reduces the dimensionality of each feature map but retains the most important information. Spatial Pooling can be based on different algorithms such as max value, average or sum of the values. The most used, and implemented in this work, is the Max Pooling: once defined the spatial neighborhood just the element from the rectified feature map within that window. Instead of taking the largest element we could also take the average (Average Pooling) or sum of all elements in that window. In practice, Max Pooling has been shown to work better.

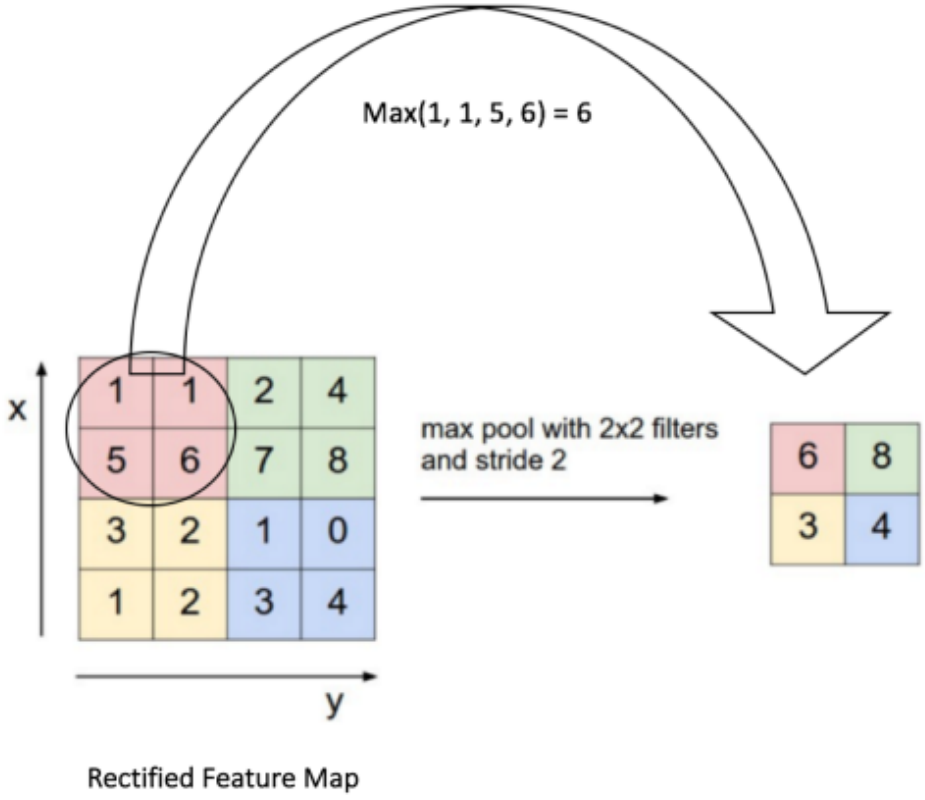


Figure 2.12: The most common downsampling operation is max, giving rise to max pooling, here shown with a stride of 2 for 2D case. Image adapted from [21]

2.6 Previous work: DeepQSM

A promising alternative to the techniques explained previously is the usage of machine learning, specially convolutional neural networks (CNN) to achieve the field-to-source problem. Implementing a convolutional deep neural network DeepQSM [38], QSMnet [48] have shown very encouraging results in term of contrast, dipole inversion, and good image reconstruction comparable to the gold-standard COSMOS technique.

On the other hand, deepQSM seems to suppress and not to consider the anisotropy effect, so the result is not perfectly correct specially when dealing with anisotropic structures. Certainly, a major source of limitation is due to the training set relied on COSMOS data that, as described before, that assumes isotropic susceptibility. Thus, extensive results can be achieved if an additional model terms accounting for anisotropy are exploited. The idea exploited in this work is to combine three different type of images, **susceptibility-weighted**, **orientation** and **FA map**, to train the neural network in order to learn the relationship between the susceptibility anisotropy and the direction of

the fibre. However, before disclosing how it can be achieved, it is important to compare the DTI and STI.

Despite both techniques are influenced by the structural anisotropy and lead to fiber tracking, the biophysical underpinnings are fundamentally different: in diffusion anisotropy, what it is detected is the hindered diffusion of water molecules due to structural barriers, so, for example, an important factor can be axonal restrictions in white mater fibers, while the susceptibility diffusion is based on the fraction of myelination of brain in white matter. This is evident from the shiverer mouse studies [26].

Mathematical differences come up because susceptibility values are relative and depend on the selection of a proper tissue reference (they can be either positive or negative), while diagonal diffusion tensor components are absolute measures of a physical constant and always positive. However, most of the studies have focused attention how STI can represent a valid alternative to DTI, that is not exactly what this work is focusing on. But, in [25] a deeper study showed results from comparing the two techniques; they demonstrated that (i) the direction of the principal eigenvalue in STI is along the fiber direction and (ii) susceptibility anisotropy is much higher sensitivity to the chemical composition of the white matter, than DTI, as it could be expected since the physical source of magnetic susceptibility. The results lead to the reason why diffusion-weighted image may provide useful indication for understanding the complexity of the anistotropic susceptibility.

A new network with a new training dataset has to be performed considering prior information provided by diffusion-weighted imaging and the attention needs to be focused on how these parameters can be fit in the network.

2.7 Aim of this work

The first challenge of the work presented in this thesis is to simulate good and appropriate synthetic data for the training set. This dataset should have a variance as big as possible in order to let the network learn the relationship between the phase image and the ground truth and must include anisotropy structures as well as information regarding orientation-dependence. Previously [38], the training set was based on random-shaped objects, randomly distributed in one 3D sample, but most of them were cubes, spheres

and rectangular cuboids so none of them included anisotropy effect. The question was, how to feed the neural network with the right information in order to understand the perturbation addicted to the orientation. The suggested solution, which has been developed in this thesis, stands in a 3-channel input neural network. The channels should perfectly approximate the parameters we need for CNN to learn, so the three inputs contain: (i) a **phase map** which is the same input given in the previous work, (ii) an **orientation map** where the angle between the main direction and the magnetic field is exploited and (iii) a **FA map** that shows where and if anisotropy structures occur, since not every brain tissue shows anisotropic effect.

The first challenge has been to implement a code that generates images containing the previous features; specifically, in order to train the neural network, high number of samples is requested by the CNN, in this case 100.000 samples per channel.

The second challenge was modify the architecture of the neural network, from 1-channel input into a 3-channel input. Most of the time has been spent in this since it was my first time approaching machine learning and Tensorflow, which is an open-source software commonly used to develop, train and perform the usage of neural networks. The third step is the training, tuning of the parameters, then testing on simulated data and validation on *in-vivo* data that the network has never seen before. Further information and more details are explored in **Section 3**.

Chapter 3

Material, Simulations and Methods

Nerve axons in the central nervous system are the main reason of susceptibility anisotropy in the brain, due to their 3D spatial orientation compared to grey matter that is isotropic. The same reason why FA is usually higher in white matter than in grey matter. This is the main reason why most of the time spent during this thesis has focused the attention how simulate nerve axons. All the methods described below have been developed using Python 3.6. When other software has been adopted, it will be explicitly reported.

3.1 Simulation: Anisotropy susceptibility effect

In order to verify and validate the 2.11 in page 22, a simulation was developed. The chosen phantom is an image whose size is 64x64x64 and it is based on two parallelepiped shaped fibers having a constant magnetic susceptibility value and located in the image space one perpendicular to the other. The forward problem has been calculated (convolution in k-space) when the magnetic field orientation changes through 14 different directions, 7 when the magnetic field rotates in one plane, 7 in an other plane. In both situation angles between the fiber and the magnetic field differ of 15° in the range [0°, 90°]. For each orientation the average phase has been measured over a ROI of 49 pixels. The results of B_0 rotating in the zy plane have been plotted in figure 3.1.

In this study the aim was to see the sinusoidal behaviour of the frequency shift when θ changes. Since the synthetic nature of the simulations and because the sample has not been placed in an external magnetic field, it was not possible to calculate the frequency

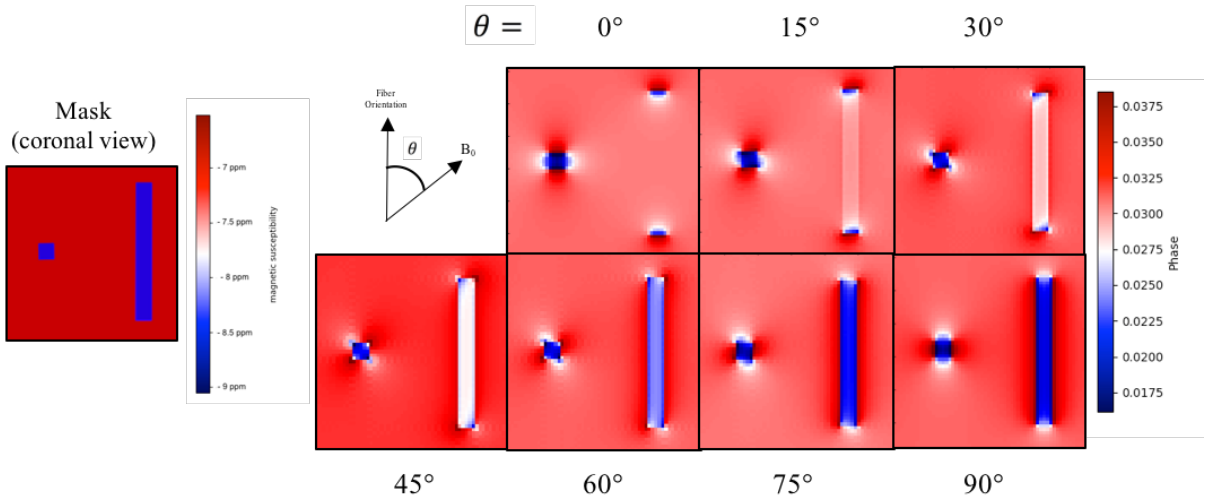


Figure 3.1: Phase perturbation around longitudinal structures when θ angle changes according to 2.10

shift. Phase data shows the same behaviour of the frequency shift and their relation is clear from 2.8 and 2.4. Figure 3.2 shows the results.

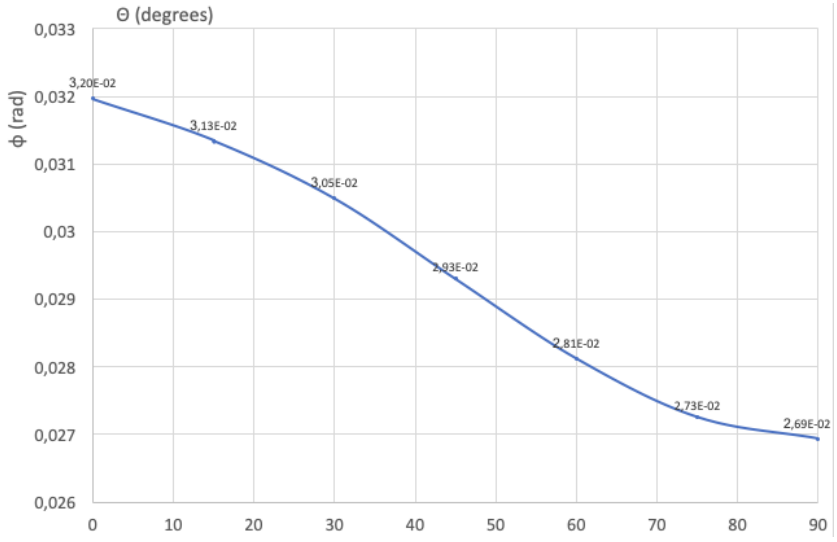


Figure 3.2: Orientation dependence averaging a ROI of 49 voxels. The behaviour should be similar to the one described in equation 2.11

3.2 Simulation: Can one assume a spherical source even in anisotropic structures?

In the literature several assumptions have been discussed in order to find the best model that suits the magnetic susceptibility source, assuming that the magnetic field perturbations could be modeled by arbitrarily-shaped objects of uniform susceptibility. First of all the Lorentz sphere has been hypothesized. Then it was demonstrated that a Lorentz cylinder seem to be more suitable [15], [47], [46]. Here it is presented a comparison between two different situations: in the first one the goal was to calculate the magnetic field changing assuming as a main source one long cylinder. In the second situation the cylinder has been considered as composed by different sources as many as the voxels within the cylinder; in this case every single source has been convolved with the dipole kernel forward problem and all the contributions have been summed together. The susceptibility values have been chosen as $\chi_{cyl} = 0.04$ and $\chi_{bg} = 0.01$ for the cylinder and the background respectively. The goal of this simulation is to prove whether the shape, resolving the inverse-problem convoluting it with the kernel is the same as a single - source distribution. In figure 3.3 shows the results of both situations and the differences.

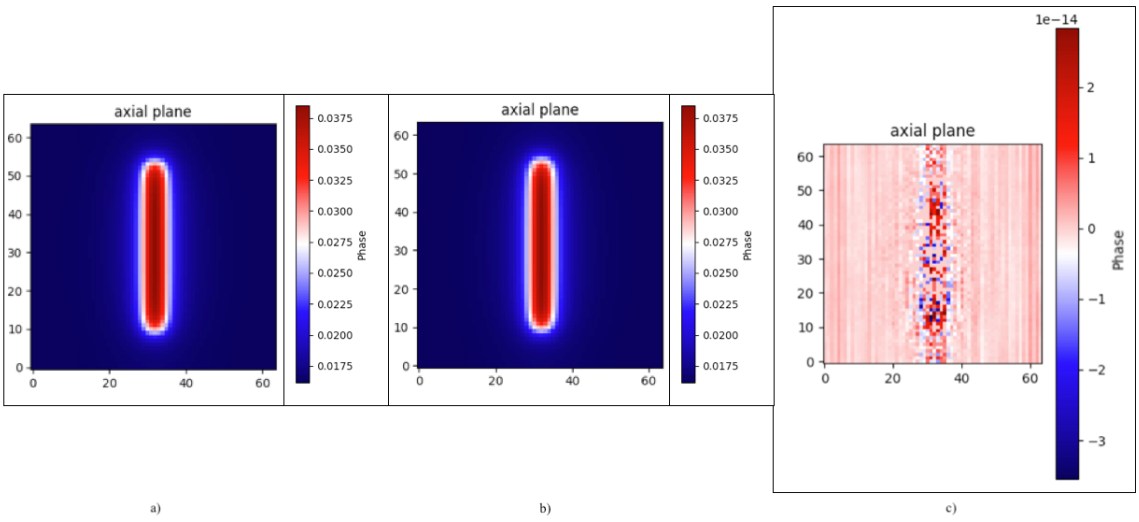


Figure 3.3: Phase image when a sphere voxel source has been considered (a), the entire cylinder as unique source (b) and the difference between the two

3.3 Simulation: Validating the training set

Dealing with simulated data has the big advantage of generating several samples and if the algorithm provides a good variability, each volume will be different enough to each other. The aim of this work is to create a model term that incorporate additional terms accounting for anisotropy of magnetic susceptibility and structural tissue anisotropy.

In order to prove that the fibers forming the training set are actually affected by anisotropic effect it has been calculated the forward and then the inverse problem for some images containing the perturbation effect. From [38] it's clear that anisotropic effect leads to an error when gold standard algorithms try to solve the inverse problem, and [38] shows that the difference between the predicted magnetic susceptibility and the ground truth is about 20 %.

In this work some tests have been performed in order to verify if the anisotropy effect actually comes into consideration when the inverse problem is performed by DeepQSM. As a result of this, one ground truth volume has been generated, composed by one sigmoid-shaped fiber. The fiber mask has been first generated using MITK [44], Medical Imaging Interaction Toolkit, in particular its toolbox, Fiberfox. MITK is an open-source software and Fiberfox is an interactive simulation tool for the generation of complex white matter tissue models. The susceptibility value assigned to the structure was 0.1 ppm. The forward problem has been performed convolving the sample with the dipole kernel and then the inverse problem has been evaluated comparing the outcome from iLSQR[1]. The forward problem, assuming COSMOS data as ground truth, has been performed in Python 3.6; iLSQR has been computer in Matlab 2016b while the deep learning approach in Python 3.6. The results of the prediction are represented in figure 3.4:

3.4 Input definition / Simulated data

In order to obtain the image dataset to feed the neural network [38], a proper algorithm has been developed. The aim is that the neural network learns the relationship between

[1] "This method uses a sparse linear equation and least-squares (LSQR)-algorithm-based method to derive an initial estimation of magnetic susceptibility, a fast quantitative susceptibility mapping method to estimate the susceptibility boundaries, and an iterative approach to estimate the susceptibility artifact from ill-conditioned k-space regions only." [5]

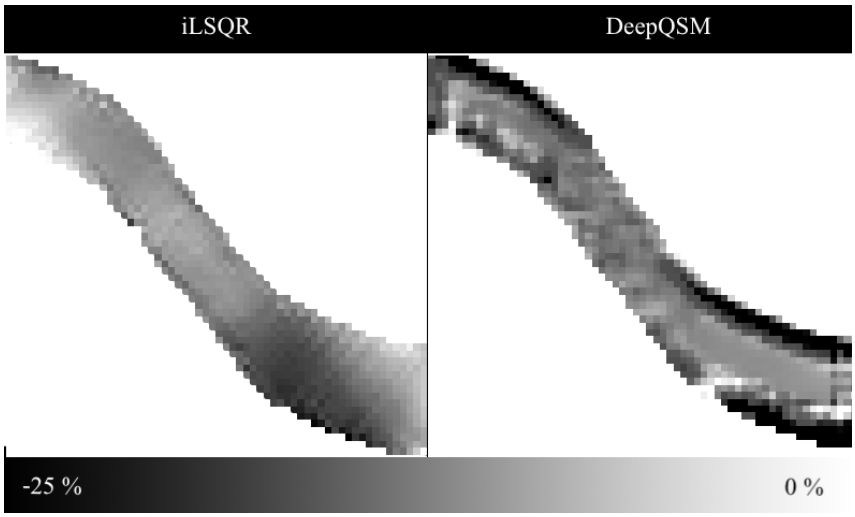


Figure 3.4: The two images show the percentage difference between the ground-truth and reconstructed Field-to-source using two different methods: iLSQR on the left and DeepQSM on the right. It is possible to see the error measurement deriving from anisotropy.

the fiber orientation to the B_0 - field and the magnetic susceptibility, so that one may be able to solve the inverse problem with *in-vivo* data.

The training dataset has been generated with image containing several fibers with different radii, orientation and position and random shapes spheres and squares. The fibers aim to simulate the phase attenuation along the main direction related to the anisotropic effect, the other random shapes have been chosen to fill up the volume simulating isotropic structures. The contribution of this thesis is to add diffusion information during network's training. DeepQSM showed good results using just random shapes object in his training set, however no diffusion information was provided. Following it is described in details the algorithm used in Python 3.6 to obtain a sample with fibers. Then a sample with random shapes as in [38] will be added. The algorithm is defined by few assumptions and conditions that the operator can choose to use or not to use.

1. An option that can be turned ON or OFF is to obtain fibers that cross each other. A certain number of crossing between the fibers can be allowed (to simulate what happens *in-vivo*);

2. Under the assumption of superposition principle, N fibers crossing result in the sum of their magnetic susceptibility values, which has been defined constant between two

fibers crossing to make a simpler mathematical model;

3. Under the same previous assumption, in the orientation map, it has been established that the orientation of multiple fibers crossing can be defined according the 2.10:

$$\chi_{cross} \sin^2(\alpha_{cross}) = \sum_{i=1}^{N} \chi_i \cdot sin^2(\alpha_i) \tag{3.1}$$

Assuming χ_i constant along crossing fibers, and χ_{cross} equal to the sum of each χ_i, with i = 1, 2,..., N, 3.1 can be simplified as following:

$$N \sin^2(\alpha_{cross}) = \sum_{i=1}^{N} sin^2(\alpha_i) \tag{3.2}$$

Inverting 3.2 allows calculation of the angle when multiple fibers cross:

$$\alpha_{cross} = \sin^{-1}\left(\sqrt{\frac{\sum_{i=1}^{N} sin^2(\alpha_i)}{N}}\right) \tag{3.3}$$

4. Phase image is obtained by the convolution of the susceptibility 3D map with the dipole kernel in z axis;

Two different strategies are reported, changing is the type of fiber which the algorithm starts with (whether it contains straight or curved fibers). Indeed, since the algorithm duplicates the starting fiber, the outcome will contain a certain amount of fibers depending on the initial input.

Straight fiber

In this case, one single straight fiber parallel to z axis has been created using MITK Fiberfox. The radius of the fiber is about 15 voxels. The aim is to duplicate the fiber several times, changing the position, rotating, flipping or scaling it until the volume condition is reached.

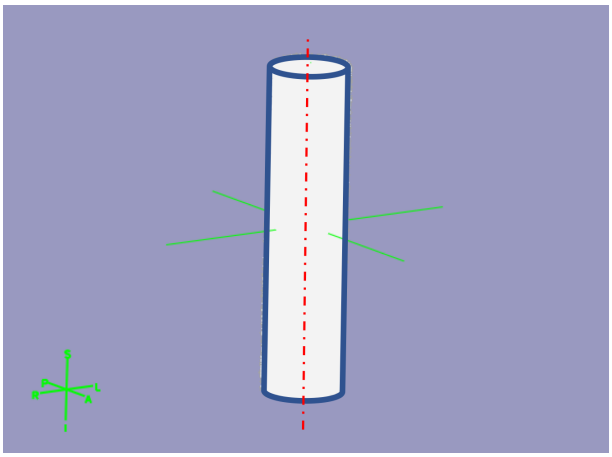


Figure 3.5: Binary mask created with MITK Fiberfox

Afterwards, in Python 3.6, the binary mask of the fiber starts to be manipulated. It is rotated about z axis randomly and the angle θ formed by z axis and the main orientation of the fiber is the one contained in the orientation map.

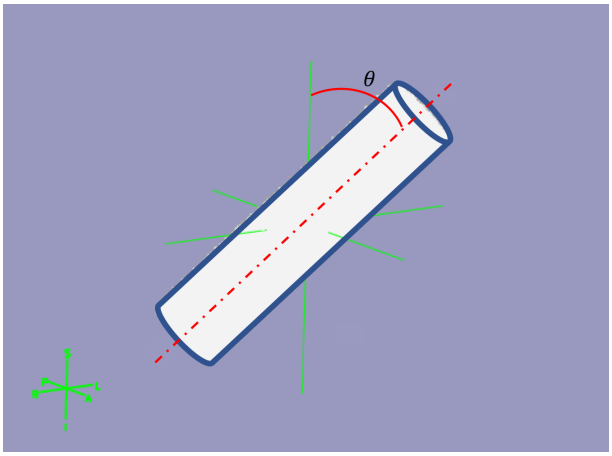
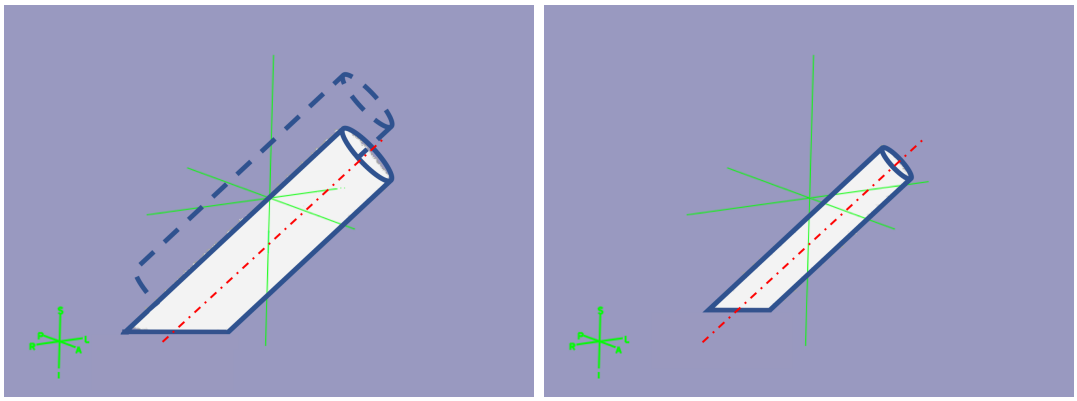


Figure 3.6: rotation

Then the mask is translated and scaled: the level of erosion depends on initial radius of the mask.



(a) Translation (b) Scaling

Advanced fiber

In this case, different fiber has been chosen as starting point, and its shapes can be randomly determined with MITK Fibermask. The advantage of using these fibers is that one can simulate structures more similar to reality and might predict better the anisotropy contribution coming from the white matter. Unfortunately this strategy is more time consuming and computationally more expensive because the clear difficulty to fill the entire volume compared to the case where the fibers are all straight.

Nonetheless, the same pipeline has been applied to this situation and the image obtained using **15000** iterations, **no overlap** in order to recreate the simplest model as possible is shown in 3.7:

The obtained volume will be called as *Sample 1*, see figure 3.7. Another volume, *Sample 2* is generated with random isotropic objects only, see figure 3.8.

The addition of the two images has been performed through voxelwise comparison: whether the voxel in the *Image 1* contains the background, it's replaced with the voxel in *Image 2*. The final result will be a configuration containing the longitudinal fibers which are not crossing the other objects, and some isotropic structures to fill the rest of the background. For training, the 3-channel input has been chosen with the following parameters: the **first channel** is the phase image coming from the convolution between the susceptibility map and the dipole kernel, exploiting the direction of the magnetic field.

What really matters in anisotropy magnetic susceptibility, is the angle between the longitudinal structure and the main magnetic field direction. For this reason the **second**

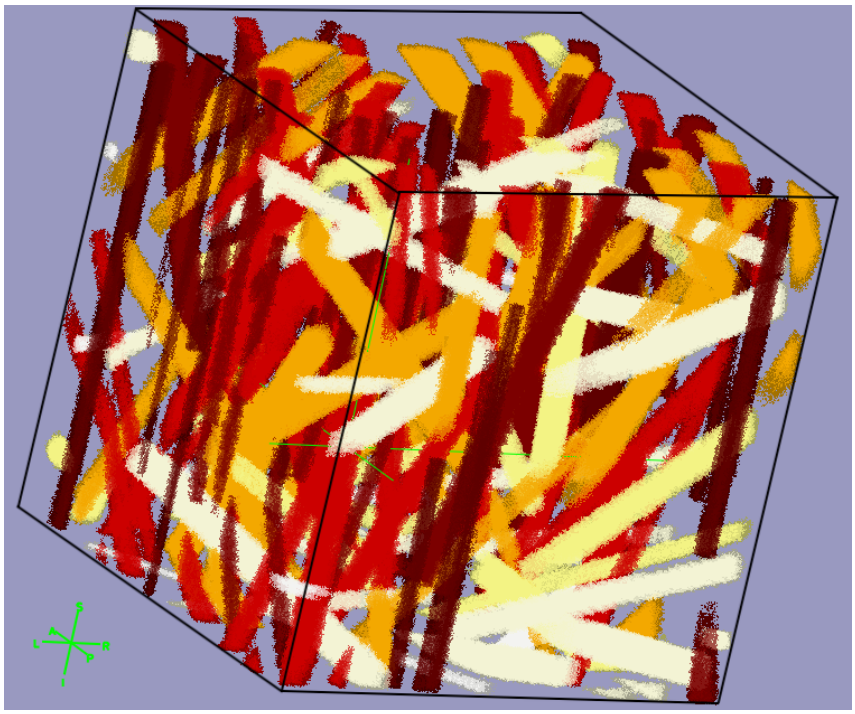


Figure 3.7: Illustration of *Image 1*: longitudinal fibers obtained after 15000 iterations, no overlap.

channel is based on the so-called orientation image that shows the $sin^2(\alpha)$ of the angle between the fibre direction and the magnetic field, with range [0, 1]. The purpose is that the CNN understands the attenuation influenced by the angle α on the magnetic susceptibility. Thus, the second channel is the sin^2 of the angle α. The *third channel* should be based on the FA map that shows if the anisotropic effect occurs. In this work, the thresholded FA map is a binary mask for the fibers since everywhere else it contains background or isotropic structures.

3.5 Training set

Once the dataset is ready, DeepQSM requires the whole training set to be considered. For this purpose, 100.000 x 3 (channels) volumes have been generated assuming this as the training set. To obtain so many images, the code has been performed using Python 3.6 on the supercomputer called Raijin, a high-performance distributed-memory cluster, procured with funding from the Australian Government and owned by NCI, National Computational Infrastructure using NVIDIA Tesla K80 GPUs. It comprises more than

Figure 3.8: Coronal section of a 3D volume sample filled with isotropic shapes only. What it is called *Image 2*.

80.000 cores in 4416 computer nodes and 300 Terabytes of main memory. Thanks to Raijin, the code has been run in parallel in multiple nodes obtaining the desired amount of samples in 23 hours.

Initially, 1000 image are generated (160 x 160 x 160 voxel size) containing the fibers in different position and the random shape to fill the background. From each sample, 100 images have been extracted (64 x 64 x 64 voxel size each) in random position, rotated or flipped in order to increase the variability. The same procedure has been done for the 3 channels. At the end, 100.000 images per channel are obtained. See Figure 3.9 for an illustration of how channel 1 was obtained during the training.

The input are scaled before feeding the CNN: any shift of the average input away from zero will bias the updates in a particular direction and thus slow down the learning rate; it is well known in machine learning that smaller ranges lead to vanishing gradients and larger ranges to exploding gradients. The scaling has been performed considering the

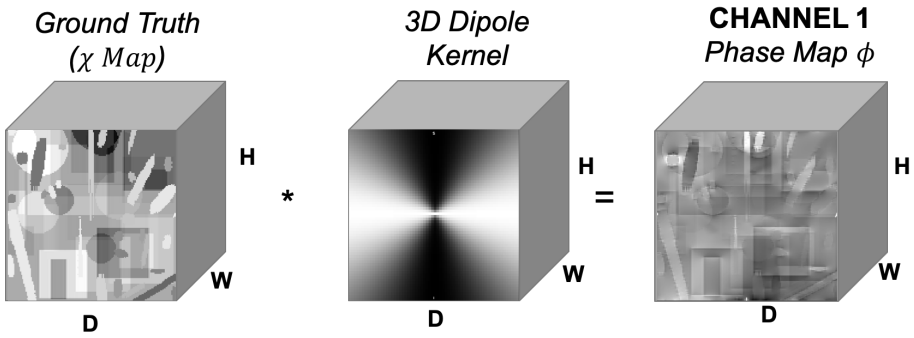


Figure 3.9: Schematic illustration that shows how channel 1 was obtained.

greatest value between the maximum and the minimum (in absolute value) and dividing the map to the nearest 10^n to that value. The final range will be [-1, 1]. The scaling has been applied just to the forward data (1^{st} channel, because the second and the third are already in that range, based on how they have been built ($\sin^2 \alpha$) and FA which *per definition* is between 0 and 1.

As first step, the training set has been generated based on **straight fibers**, **no overlapping** among the fibers to make the model simple as possible. The amount of volume occupied by the fibers is around **20 %**.

Finally the training has been performed in Raijin. Three different training have been performed: 3 channel input, 2 channel (phase image and orientation) and 1 channel (phase image only). To optimize the weights of the network, 'ADAM' (ADAptive Moment estimation) optimizer has been used. As explained in [19], it is an optimization algorithm that can be used instead of the classical stochastic gradient descent procedure to update network weights. Instead of adapting the parameter learning rates based on the average first moment (the mean), ADAM also makes use of the average of the second moments of the gradients (the uncentered variance). Specifically, the algorithm calculates an exponential moving average of the gradient and the squared gradient, and the parameters β_1 and β_2 control the decay rates of these moving averages. ADAM is a popular algorithm in the field of deep learning because it achieves good results quickly In this work, the following configurations have been chosen: initial learning rate = 0.001, β_1 = 0.9, β_2 = 0.99, as suggested by [19]. The cost function is the mean squared error between the reconstruction from U-NET and the label data. The training has been performed using Tensorflow v1.8, an open-source software library for machine learning applications such

as neural networks, developed by Google team.

3.6 DeepQSM Architecture

In order to achieve high accuracy with the constraints of a three-dimensional model, a fully convolutional neural network named as DeepQSM [39], [7] was chosen. This architecture allows challenging biomedical 3D images segmentation with an accuracy close to human performance. In this work, the convolutional layer is performed by a 3 x 3 x 3 size kernel, after padding the image so the size is constant, followed by a Rectified Linear Unit (ReLU)[2] which is a non-linear operation and strides of one dimension each. Figure 3.10 shows this function.

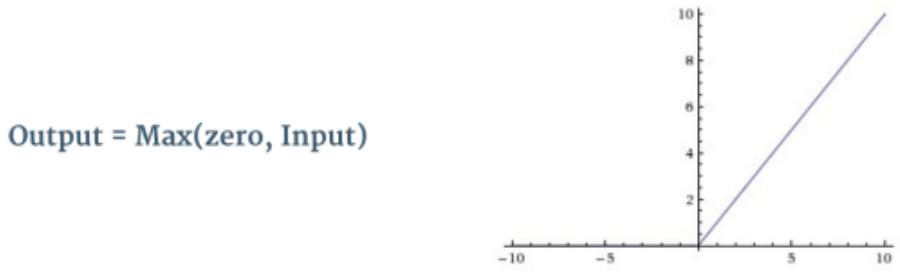


Figure 3.10: Plot of the rectifier ReLU.

The number of the kernel chosen is 16, similar to the resulting features at the very first level. It follows a second convolution but in this case the number of the features remains constant. After each convolutional layer, the *dropout* function is applied. Dropout consists in randomly setting a fraction rate of input units to 0 at each update during training time, which helps prevent overfitting. Then, the downsampling is performed by 2 x 2 x 2 max pooling with strides of two in each dimension so the image size is halved. This scheme is repeated 4 times until (the bottom part of the image) the image size is 16 time smaller then the input size and the number of the features is 256 in total. In the second part, the upsampling is performed (or expanding phase) where each layer consists of an upsampling of 2 x 2 x 2 by strides of two in each dimension, followed by two 3 x 3 x 3 convolutions each followed by a ReLu.

[2]ReLU is an element wise operation (applied per pixel) and replaces all negative pixel values in the feature map by zero. The purpose of ReLU is to introduce non-linearity, since most of the real-world data that ConvNet has to learn would be non-linear. Note that convolution is a linear operation element wise matrix multiplication and addition, so the non-linearity is introduced only by ReLU.

In the last layer a 1 x 1 x 1 convolution reduces the number of output channels to the number of labels which is 1 in our case.

The input to the network is a 4D volume (64 x 64 x 64 x 3) where 3 stands for the channel layers with 3 layers. The output in the final layer is a single 3D volume image (64 x 64 x 64) where the inverse problem has been be performed.

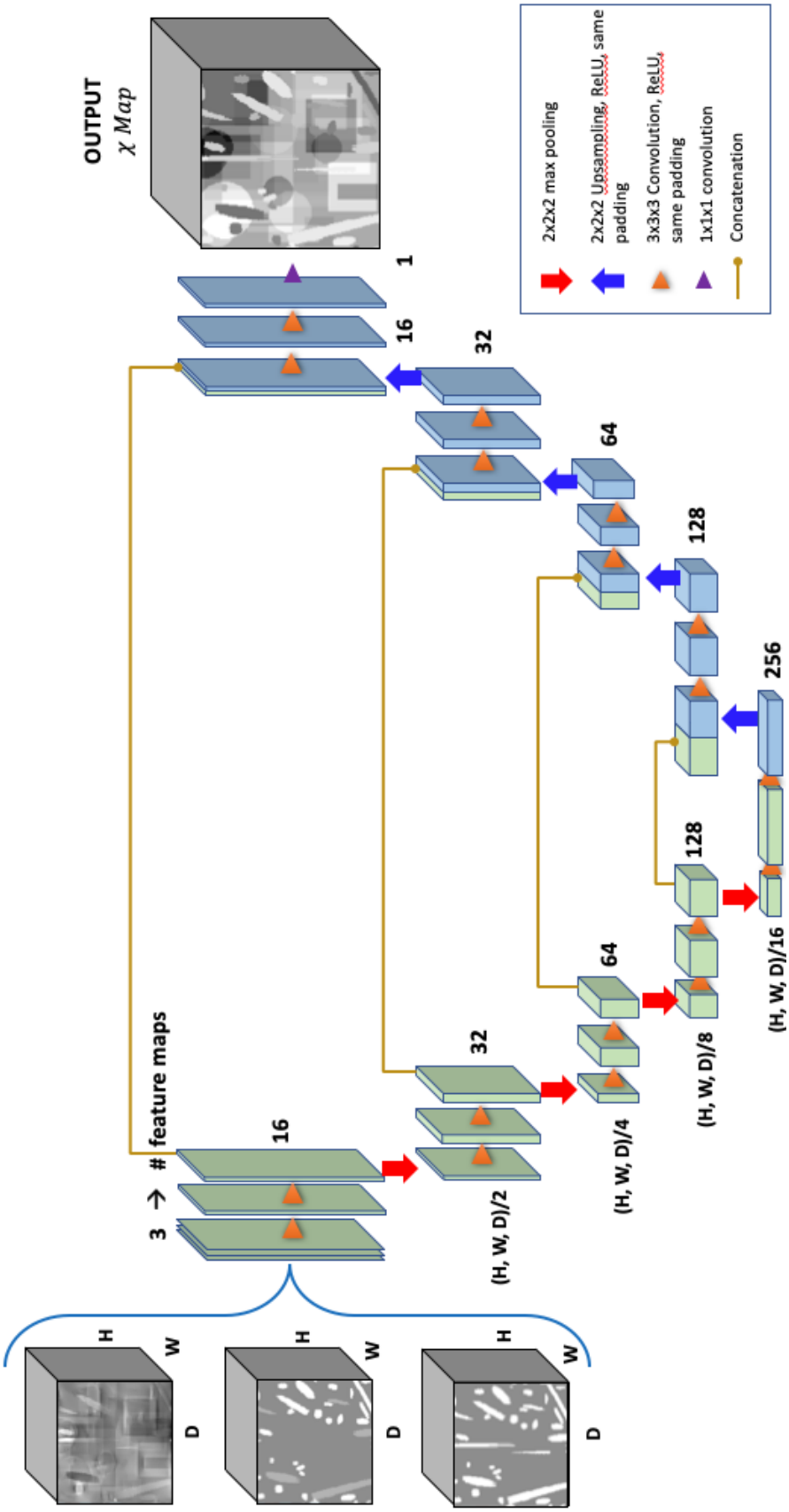


Figure 3.11: Architecture of DeepQSM

3.7 Training in Tensorflow

As discussed in **Section 3.5**, the training phase has been performed using a 3 channel input neural network. However, 1 channel and 2 channels input has been tested. In any of these cases, the expected output is the χ map where the inverse problem is reverted. Each training has been performed using Tensorflow library in Raijin. Each training required 15 GB of memory and time depending on the number of epochs chosen. In Tensorflow, there are some parameters that need to be set:

1. **Batch size**: this corresponds to the number of samples that the network can be fed in one single step. It is strongly influenced by the memory usage of the computer it is run with. In our case, a batch size of 40 samples has been chosen and it required a memory size of 15 GB.

2. **Epochs**: they refers to the number of times the neural network will be trained by one full dataset. In our case this was 4 because of time constraints.

3. **Dataset Shuffle ON**: this randomly shuffles the elements of the dataset every epoch before feeding the network. It helps avoiding overfitting.

4. **Checkpoint step save**: this indicates how many steps (each step correspond to training a single batch size) it is required to save the checkpoints of the network during the training. In our case it has been set as 500 steps.

It is important to know how many steps the training need to undergo before finishing, and it is calculated as follows:

$$\#steps = \frac{\#samples \cdot epochs}{batch\ size} \tag{3.4}$$

In this work case, the total number of steps is 10,000. To see the performance of the training, Tensorflow is able to save the so-called checkpoints parameters so that the behaviour of the loss function and the weights can be monitored. The loss function tested in this work are Absolute Difference (L1) and Mean Squared Error (L2). Both functions have been tested on the same network to see whether there is any difference in the training. Dropout rate has been chosen relatively low for each experiments (0.05-0.15).

3.8 Validating the network

During the training the **validation set** will measure the performance of the network in order to understand when and if the overfit will occur. This dataset is composed of simulated data coming from the same algorithm described before even though they represent data the network has never seen previously (unseen data). The accuracy between the predicted image from the validation set and the ground truth has been performed using Mean Squared Error during the training.

One more validation will be performed including *in-vivo* data. This process will check the performance of the neural network in terms of solving the inverse problem, different experiments have been performed. This test whether the neural network is able to solve the inverse problem generated by the dipole kernel in brain images that it has never seen before. For this purpose COSMOS data was chosen (first channel) and Diffusion data from the other two. The matter was processed from 12 orientations (see **Section 2.3**) and this represents the Ground Truth. The phase map is originated convolving the ground truth with the dipole kernel. The orientation information comes from DWI imaging, related to the same patient. Both techniques were acquired using 7T Siemens scanner and further details regarding the acquisition are showed in Table 3.1.

3.9 Preprocessing *in-vivo* MRI data

The preprocessing step is fundamental in order to obtain consistent and denoised data to test DeepQSM's performance. Another important practical reason is that raw COSMOS or DWI data can't be used as input channels of the network. 12 orientation has to be processed in order to solve the inverse-problem according to COSMOS pipeline, the background field has to be removed and the multiple orientations have to be registered to

Table 3.1: Summary from MRI acquisition used for validation. Both images was acquired from a young healthy 24 year old male patient, using 7T Siemens scan in the Melbourne Brain Centre Imaging Unit.

MRI acquisition parameters						
	Acquisition	Slice thickness (mm)	TE (ms)	TR (ms)	Resolution (mm)	b-values
COSMOS	GRE	0.6	4.8	18	0.6 (iso)	/
Diffusion	EPI	1.24	72.4	7000	1.24 (iso)	104

a reference one. Finally the field-to-source problem can be reconstructed. In Figure 3.12 all the steps required.

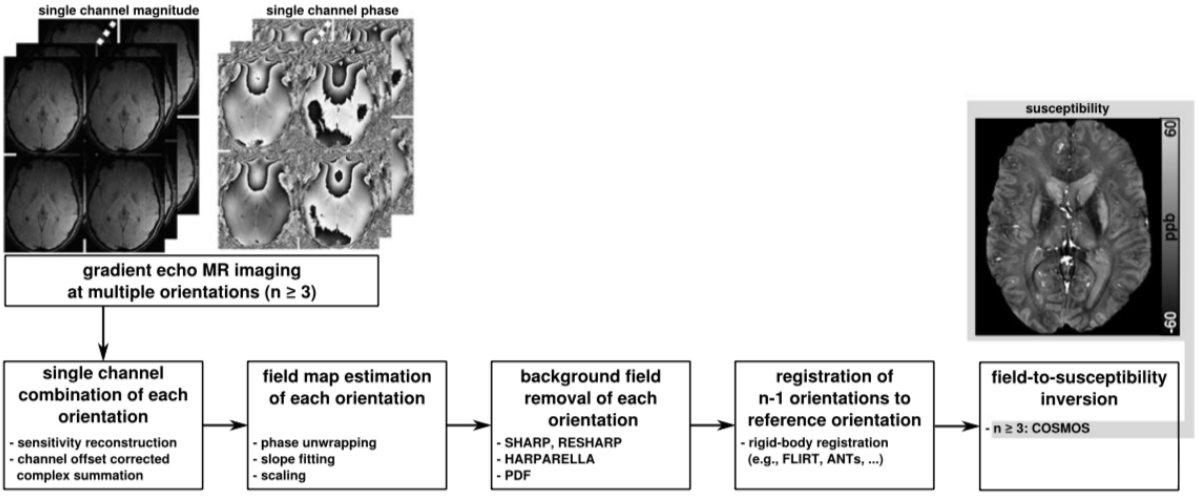


Figure 3.12: Processing steps to convert gradient echo (GRE) phase images into magnetic susceptibility for multiple orientation susceptibility computation (COSMOS). Image adapted from [8].

On the other side, DWI image has to be processed in order to extract an the orientation angle of each voxel respect to the magnetic field and FA map. After denoising, inhomogeneity field estimation must be performed: EPI MRI acquisition suffers of Phase Encoding distortion. Two PE are required to compensate the distortion: Anterior to Posterior (A2P) and Posterior to Anterior (P2A). MRtrix preprocessing techniques allows to compensate such distortion and obtain the expected proportion of the acquisition. Afterwards the mask has to be computer, then FA and orientation map are calculated. Nevertheless both images are in the diffusion space. All the three channels must be in the same space, therefore rigid-body-registration to COSMOS space is required. In this case linear followed by non-linear registration has been performed using nearest neighbour (NN) interpolation. In figure 3.12 the preprocessing steps are listed.

The orientation map presents the angle of the voxel that contains only structure with FA greater than 0.5. In this case, FA has been thresholded to 1 when. In other words, FA contains a sort of binary mask highlighting whether a voxel contain an anisotropic structure, and orientation map has $\sin^2(\theta)$ of such structures.

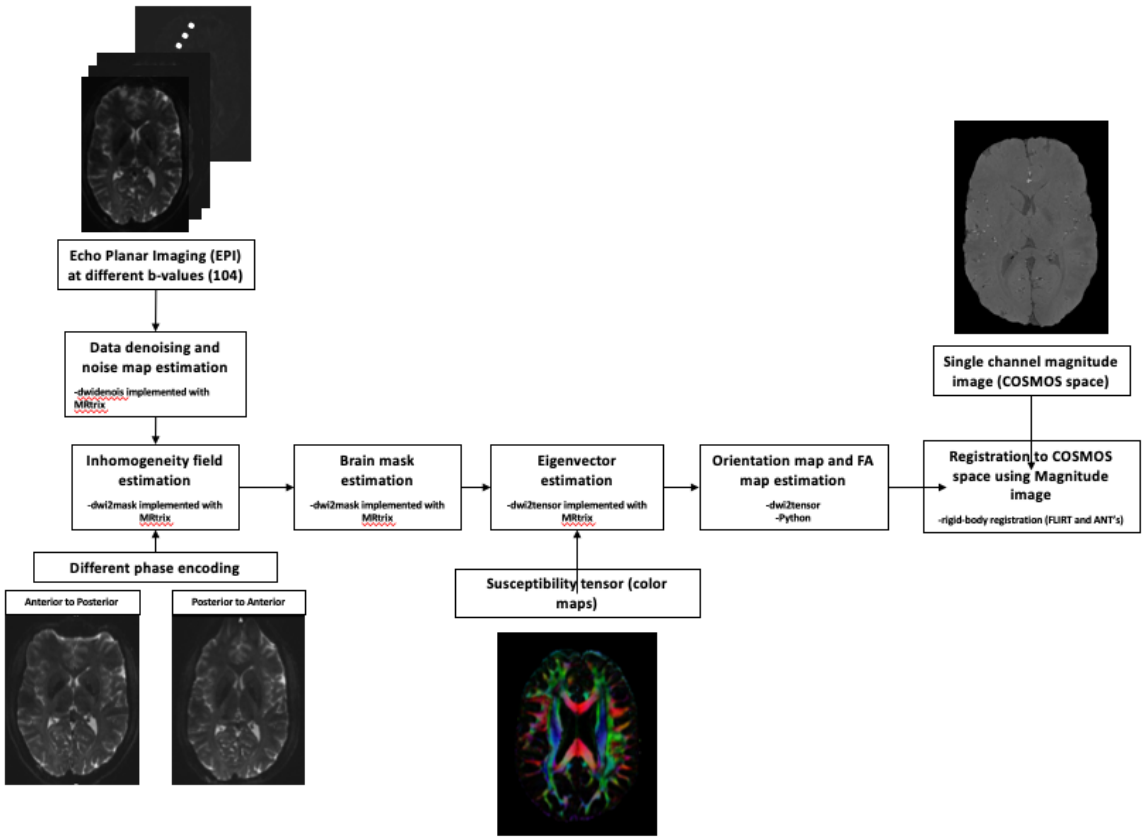


Figure 3.13: Processing steps required to obtain the second and third channel for validation phase. to convert gradient echo (GRE) phase images into magnetic susceptibility for multiple orientation susceptibility computation (COSMOS). Image adapted from [8].

Chapter 4

Results

As explained previously in Chapter 3, many training steps have been performed in order to tune different network parameters. The aim is to see whether they fit better or less better the inverse problem. For this reason different situations have been explored. Listed below there are the main results coming from training, evaluation and test process. 1-channel, 2-channels and 3-channels input have been tested using L1 or L2 as loss function. Accuracy, training and evaluation's loss function have been represented in the following. However Figure 4.1 shows an evolution of the training (it refers to training in Figure 4.4, even though other cases studied showed a similar behaviour).

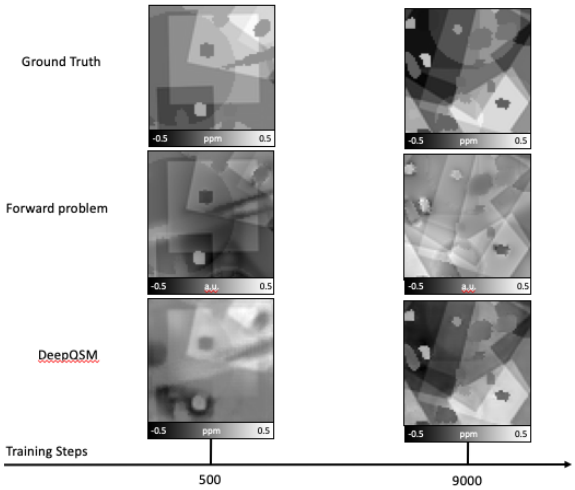


Figure 4.1: Typical prediction of the validation set during the training process.

4.1 Input definition

Figure 4.2 presents how the forward problem is applied on *in-vivo* brain images used to perform the test process. The second row of the same image shows the 1^{st} channel of the input.

The following results are listed classifying them per channel-configuration. Each of them will include results from training, evaluation and test. Figure 4.3 shows an axial

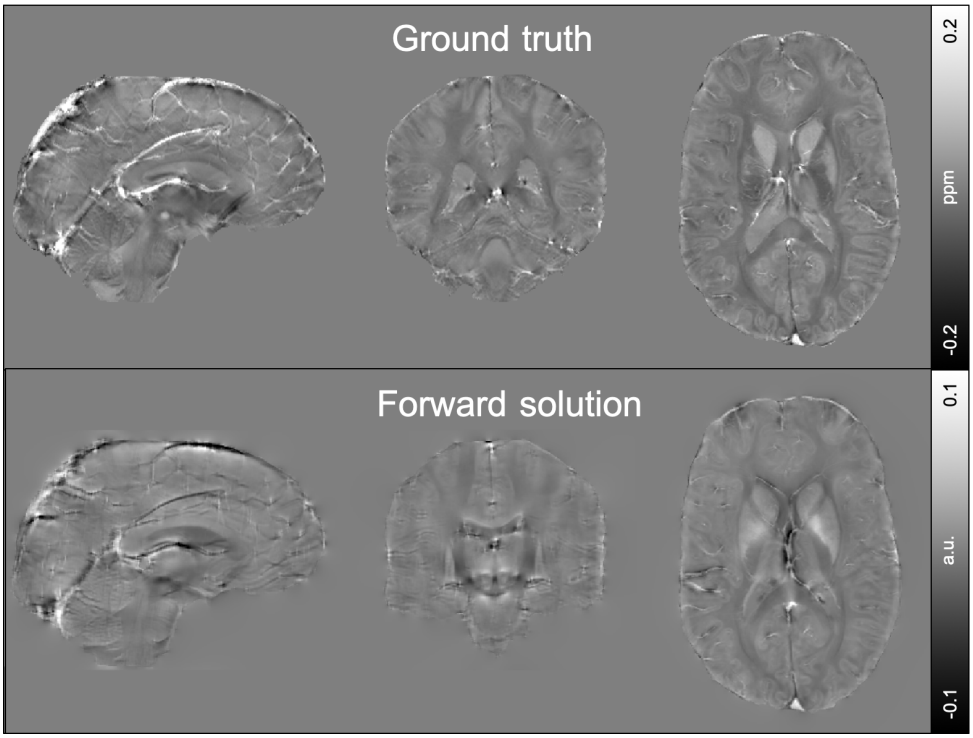


Figure 4.2: Ground truth representation (top) and forward solution (bottom). The first row shows a sagittal, coronal and axial slice of the COSMOS data from 24 year old healthy male patient. The second row presents the same representation but related to the forward problem. Both images are used during test process. The forward solution is the first channel of the network.

slice of the three channels (from the left forward solution, orientation map and FA). II and III channels have been thresholded as discussed in section 3.9. The threshold has been made to underline the structures where the anisotropy effects cause such perturbation that we want to be corrected by the network.

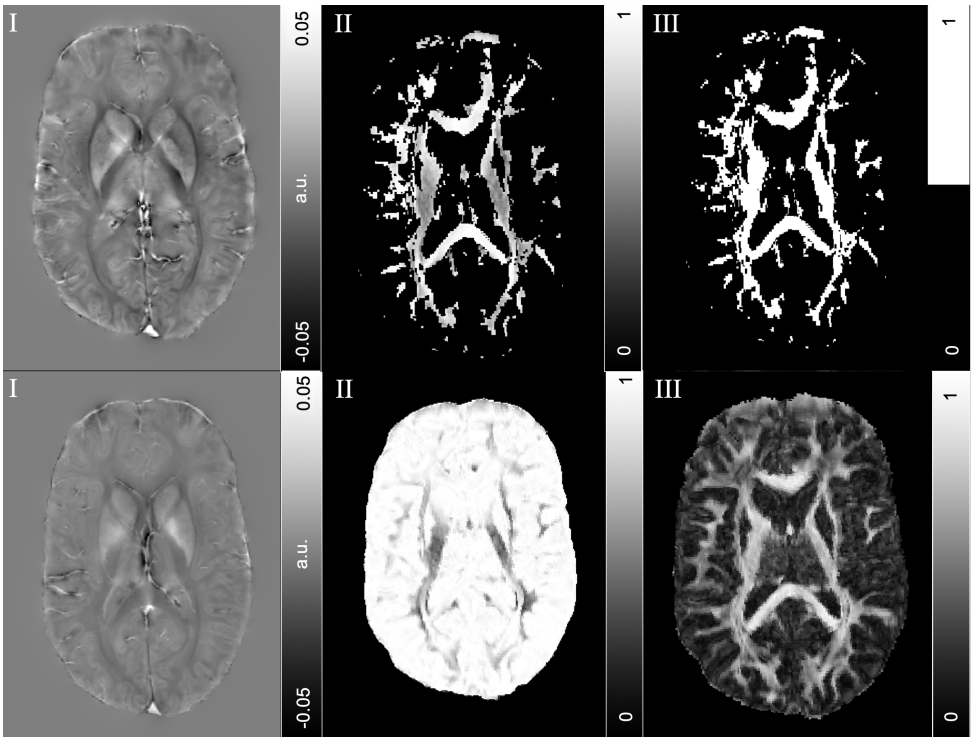


Figure 4.3: Comparison between two inputs, when FA is thresholded (top) and when it is not (bottom). Both rows shows an axial slice of each channel during the test process.

4.2 3 channels

Figure 4.4 and 4.5 presents two training processes associated to 3-channel input neural network. In order to evaluate the training process, each graph reports the accuracy and loss function, for both training and validation set. In both case the dropout rate was set equal to 0.15 and 4 epochs were chosen. L1 and L2 loss function have been evaluated in two different scenarios.

Comparing the two results, it can be seen that the loss function seems quite different between Figure 4.4 and 4.5. However, these results do not indicate how good is the prediciton on real *in-vivo* data. The results from the prediction can be compared in Figure 4.6 and 4.7. The phase map has been obtained convolving the COSMOS/ground truth brain image, as summarised in Figure 4.2.

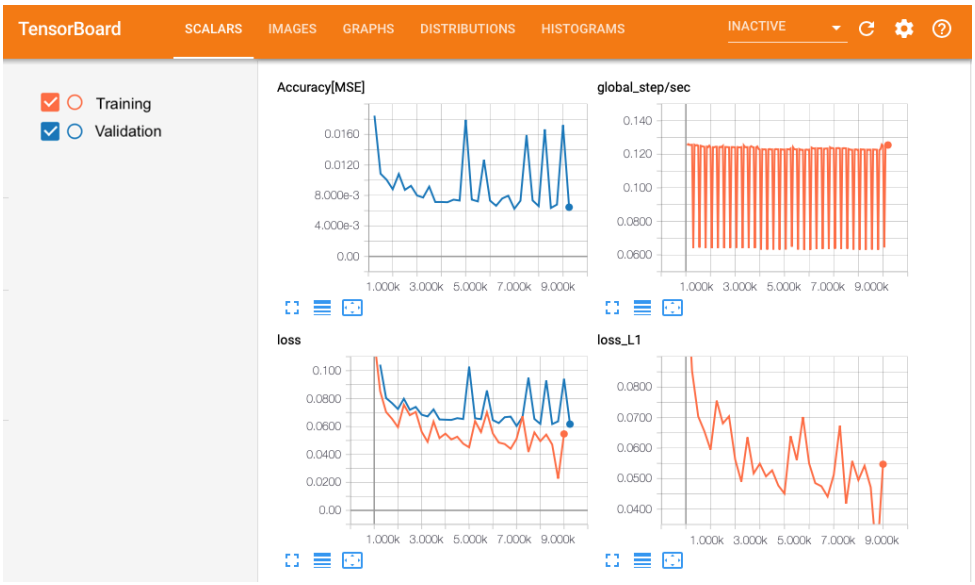


Figure 4.4: Training process using 3 channels input, 4 epochs and absolute difference (L1) as loss function.

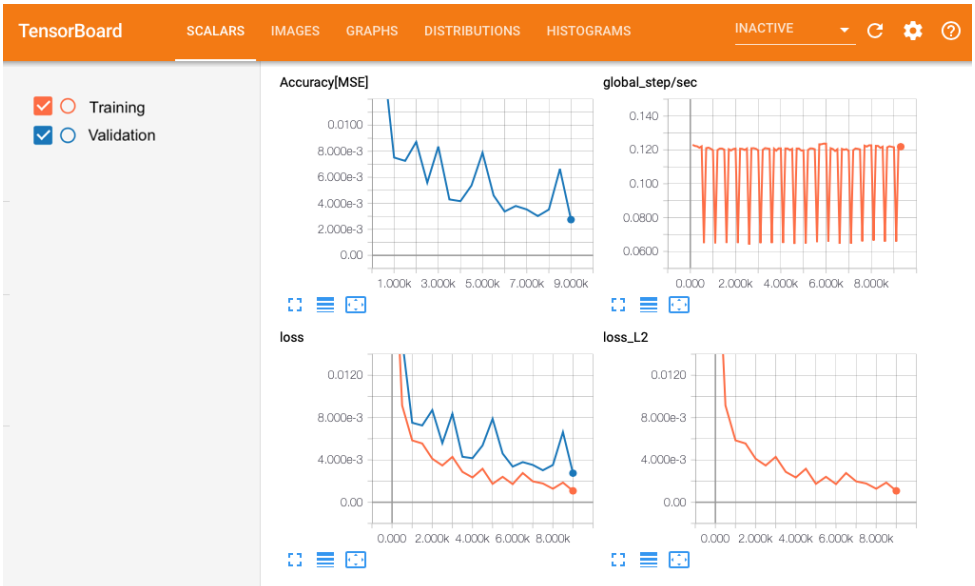


Figure 4.5: Training process using 3 channels input, 4 epochs and mean squared error (L2) as loss function.

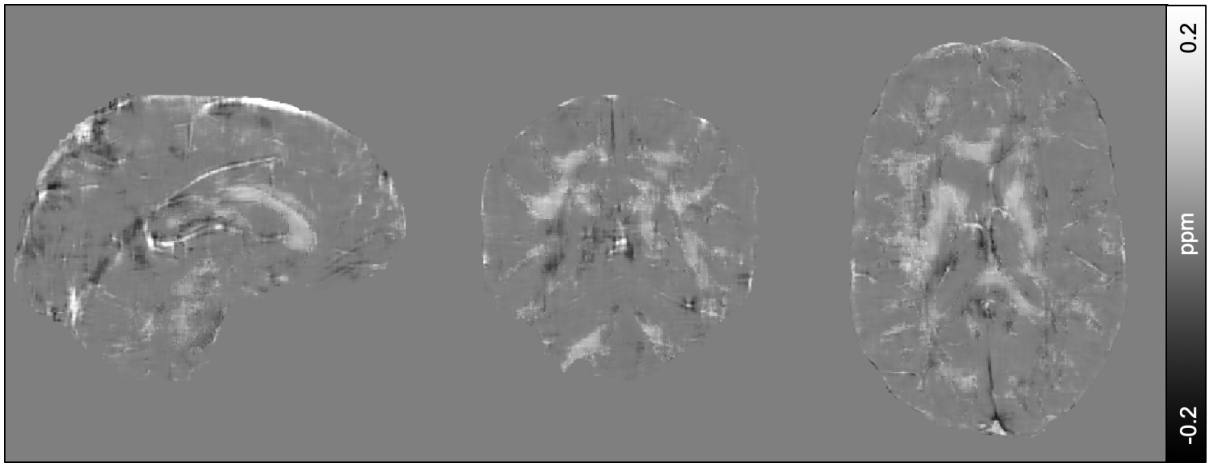


Figure 4.6: Brain image validation for 3-channel input and loss function L1.

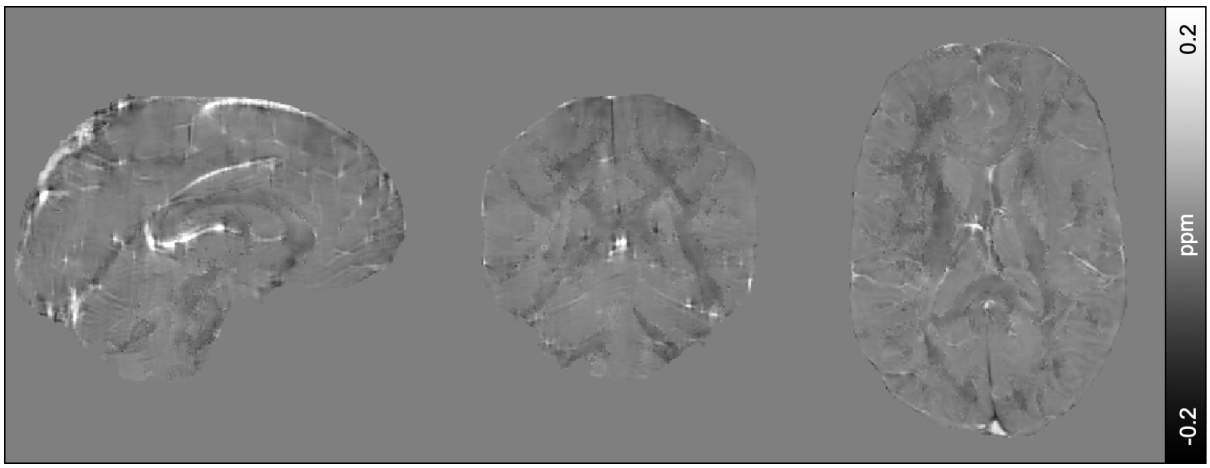


Figure 4.7: Brain image validation for 3-channel input and loss function L2.

What is interesting in this data is that both predictions presents a sort of artifact effects in that point.

4.3 2 channels

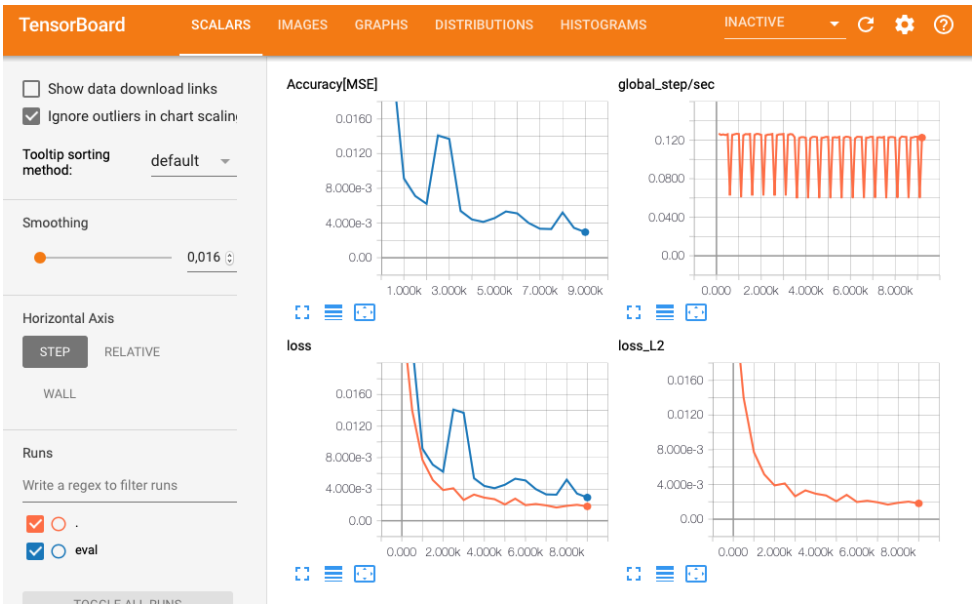


Figure 4.8: Training process using 2 channels input, 4 epochs and absolute difference as loss function (L1).

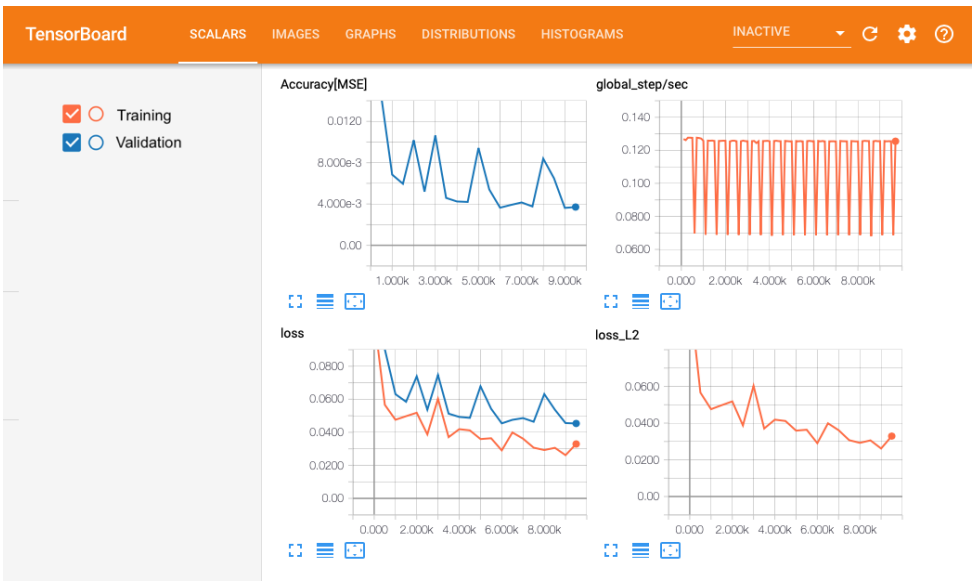


Figure 4.9: Training process using 2 channels input, 4 epochs and mean squared error as loss function (L2).

Differently from the previous results, Figure 4.8 and 4.9 shows the training process when 2-channels input is chosen. The first channel is always the phase map, the second channel is an orientation map. Figure 4.8 and 4.9 summaries the training parameters.

Figure 4.10 and 4.11 are the results from the prediction of the previous training process when *in-vivo* brain image is the input.

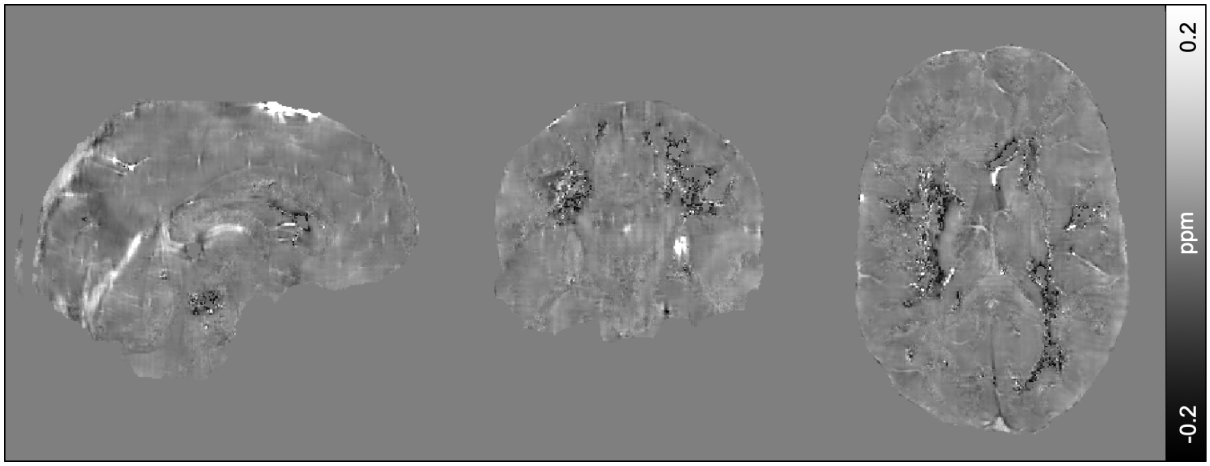


Figure 4.10: Brain image validation for 2-channel input and loss function L1.

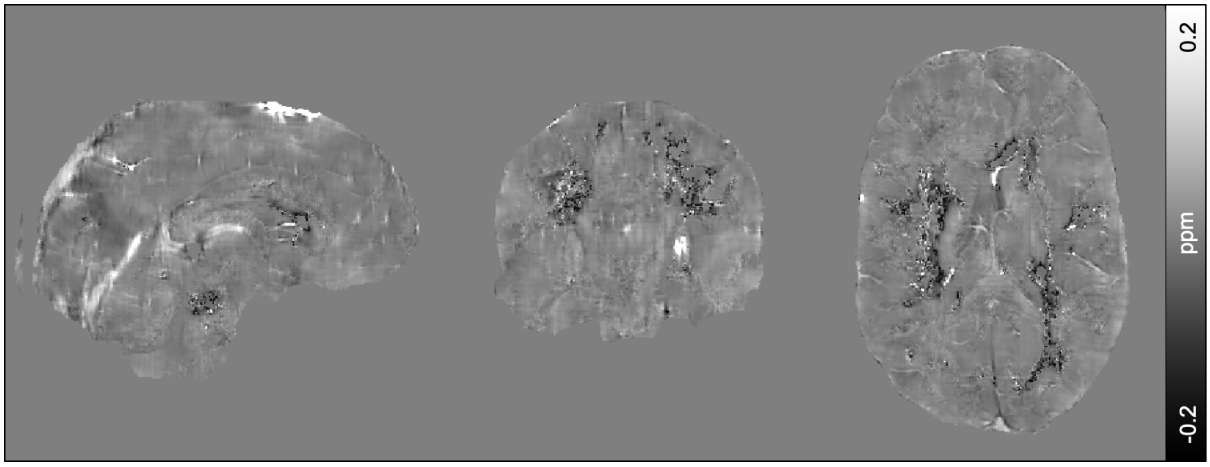


Figure 4.11: Brain image validation for 2-channel input and loss function L1.

4.4 1 channel

1 channel input DeepQSM is similar to the one presented in [38], the only difference is the training set used during the training. Figure 4.12 and 4.13 shows the training process using L1 and L2 as loss function respectively.

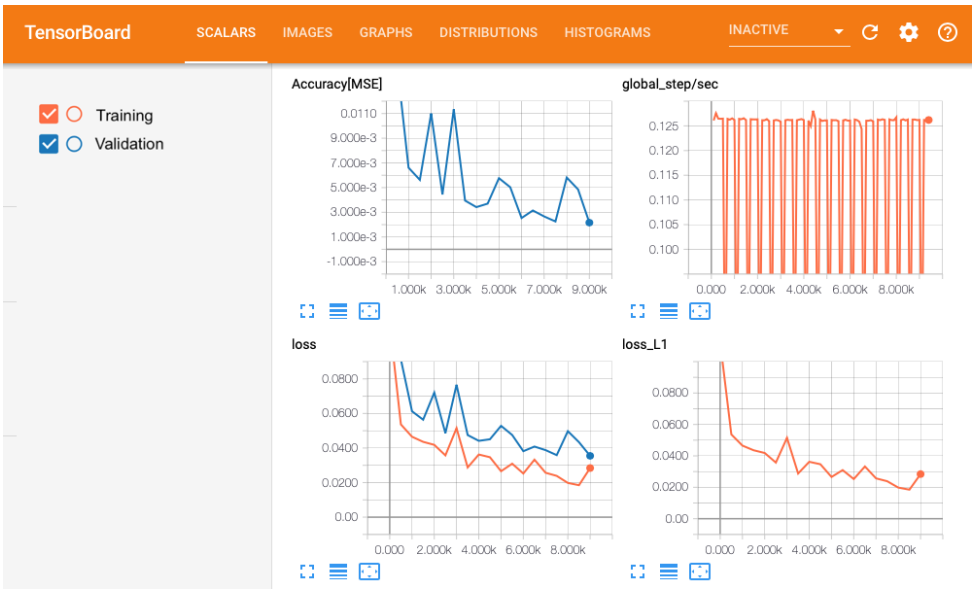


Figure 4.12: Traning process using 1 input channel, 4 epochs and absolute difference as loss function (L1).

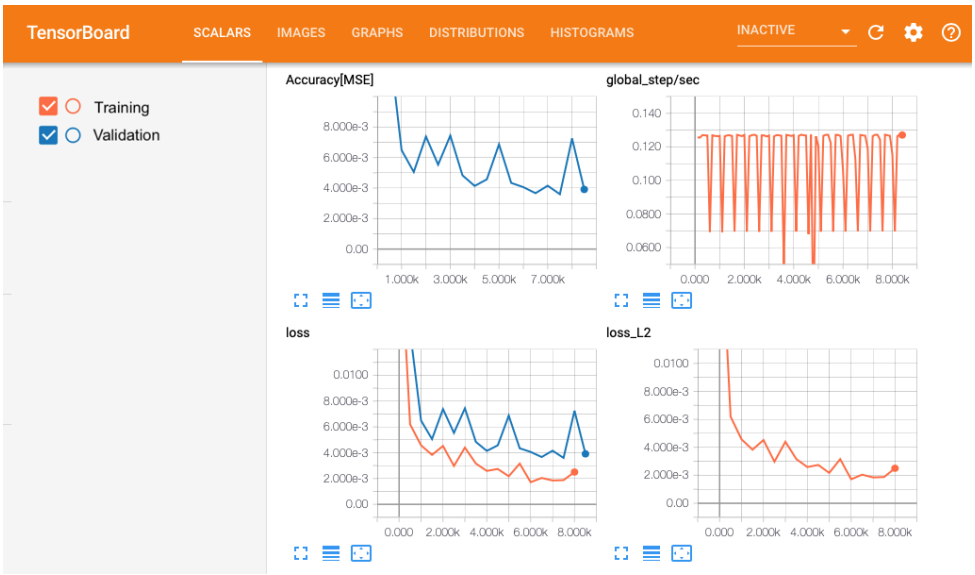


Figure 4.13: Traning process using 1 input channel, 4 epochs and mean squared error as loss function (L2).

Figure 4.14 and 4.15 shows the validation outcome of the previous network architecture.

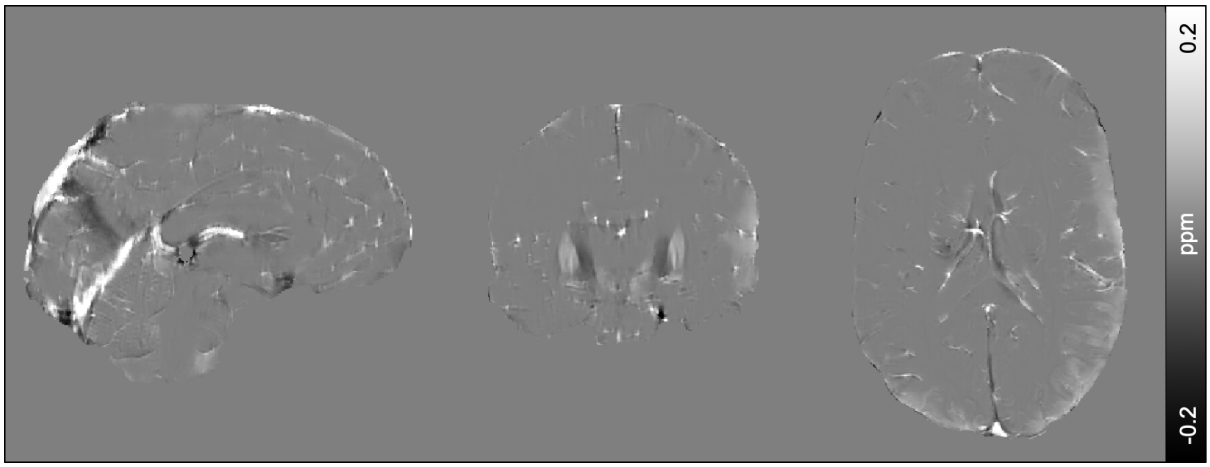


Figure 4.14: Brain image validation for 1 channel input and loss function L1.

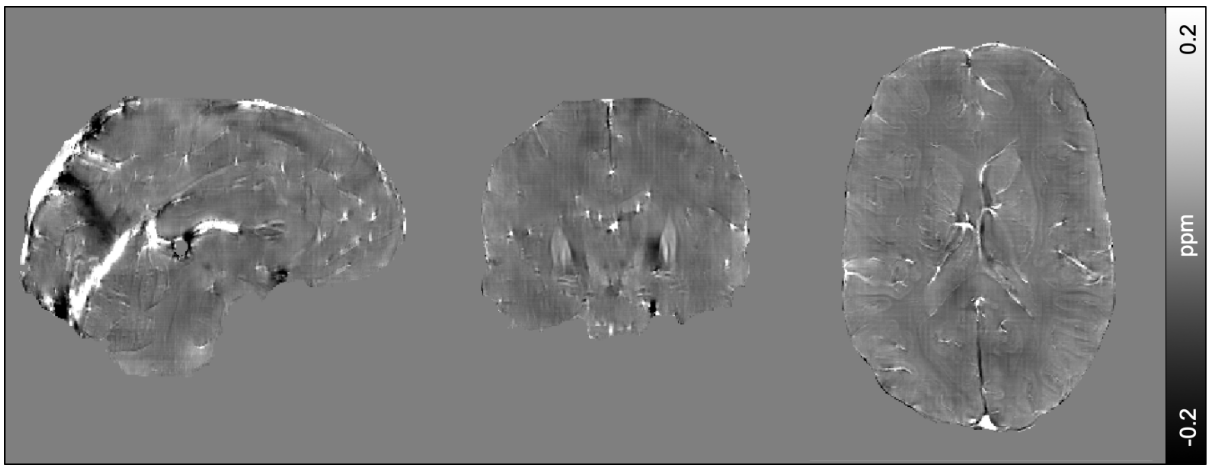


Figure 4.15: Brain image validation for 1 channel input and loss function L2.

The previous predictions were performed using a threshold input (top image in figure 4.3). Using the image at the bottom in figure 4.3, the results from the prediction are as follows:

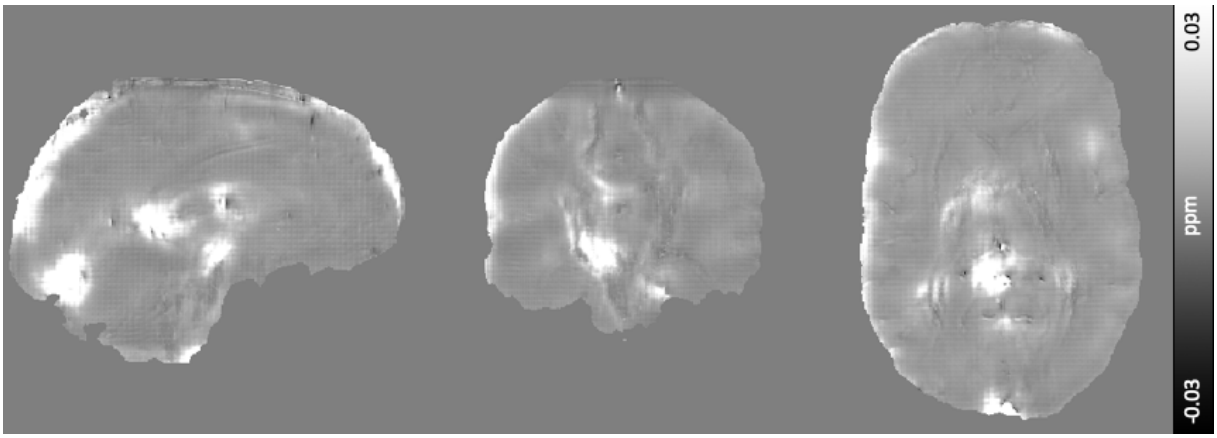


Figure 4.16: Brain image validation for 2 channels input and loss function L1 (no threshold input).

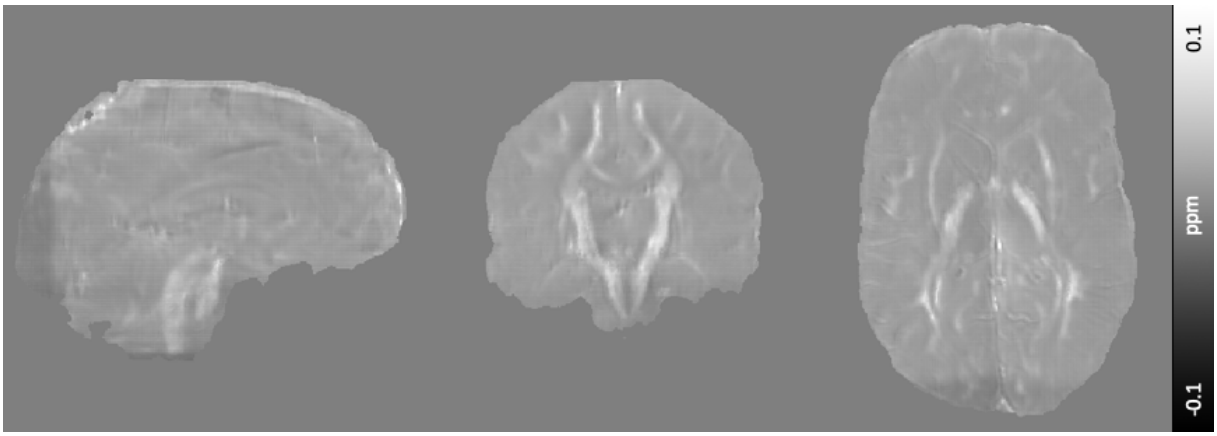


Figure 4.17: Brain image validation for 2 channels input and loss function L2 (no threshold input).

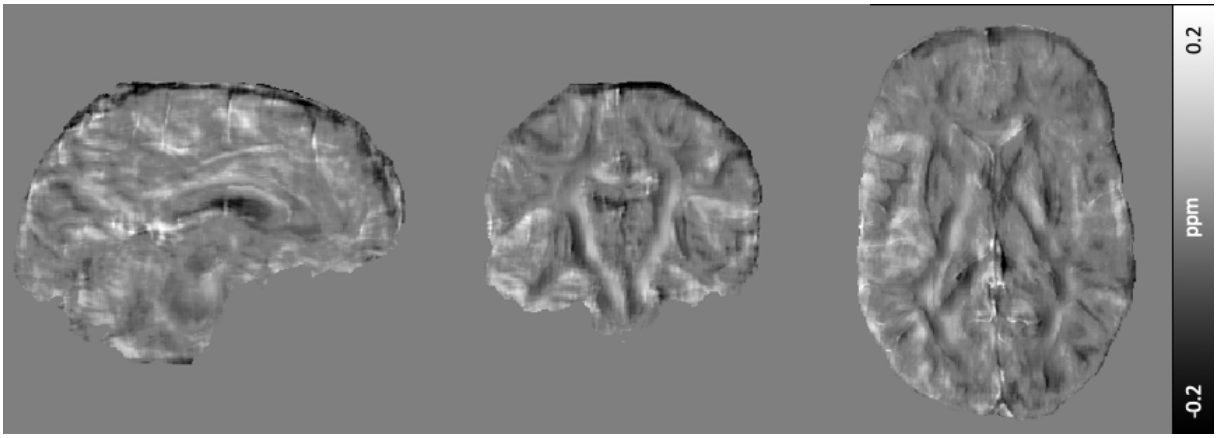


Figure 4.18: Brain image validation for 3 channels input and loss function L1 (no threshold input).

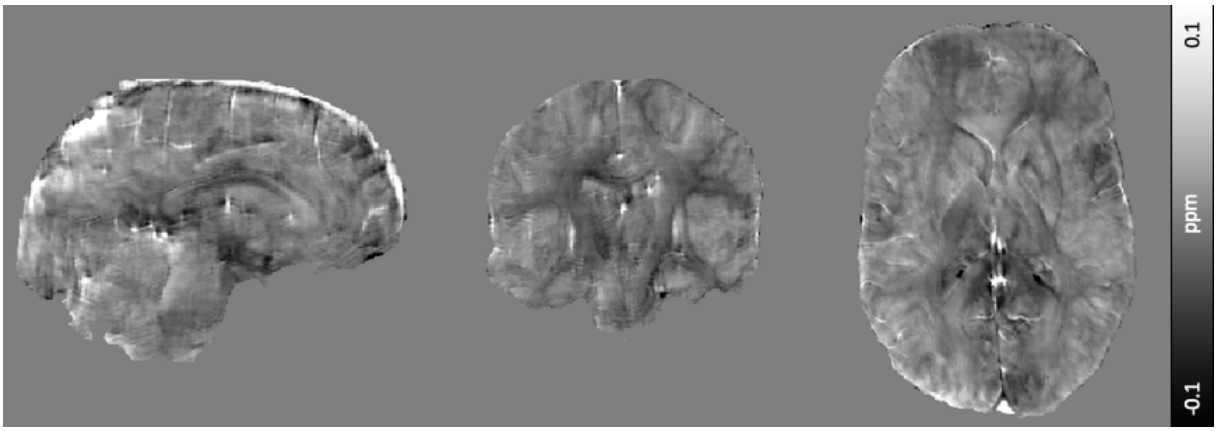


Figure 4.19: Brain image validation for 3 channel input and loss function L2 (no threshold input).

4.5 Comparison

For each test prediction Mean Squared Error (MSE) and Structural Similarity(SSIM) [1] have been calculated. Figure 4.20 and 4.21 respectively, shows

[1]SSIM is computed from three distinct terms : luminance (mean), contrast (variance) and structure (correlation), computed through statistical parameters calculated calculated in a sliding Gaussian weighted window usually about 11 pixels wide. The three terms are then multiplied together at each pixel location to produce a SSIM map which is subsequently uniformly pooled to obtain a single SSIM index.

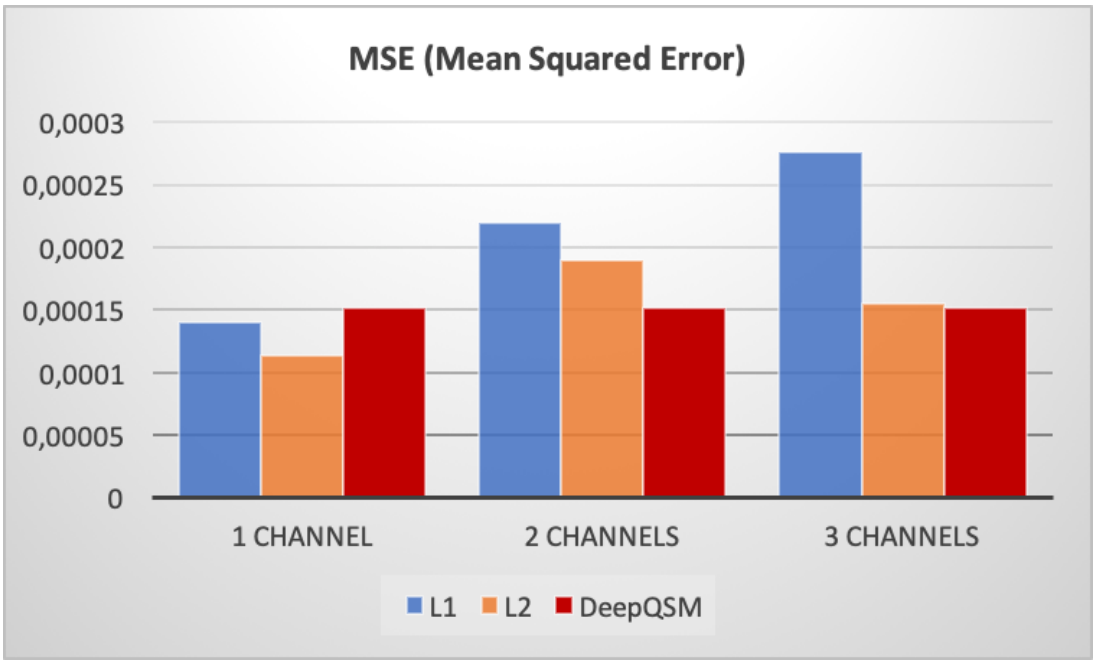


Figure 4.20: MSE computed between on the predicted brain image and the lable used during the validation process. The blu bar shows the outcome the network using L1 as loss function, the orange bar shows L2 loss function based network and the red bar shows the MSE computed using DeepQSM from [38].

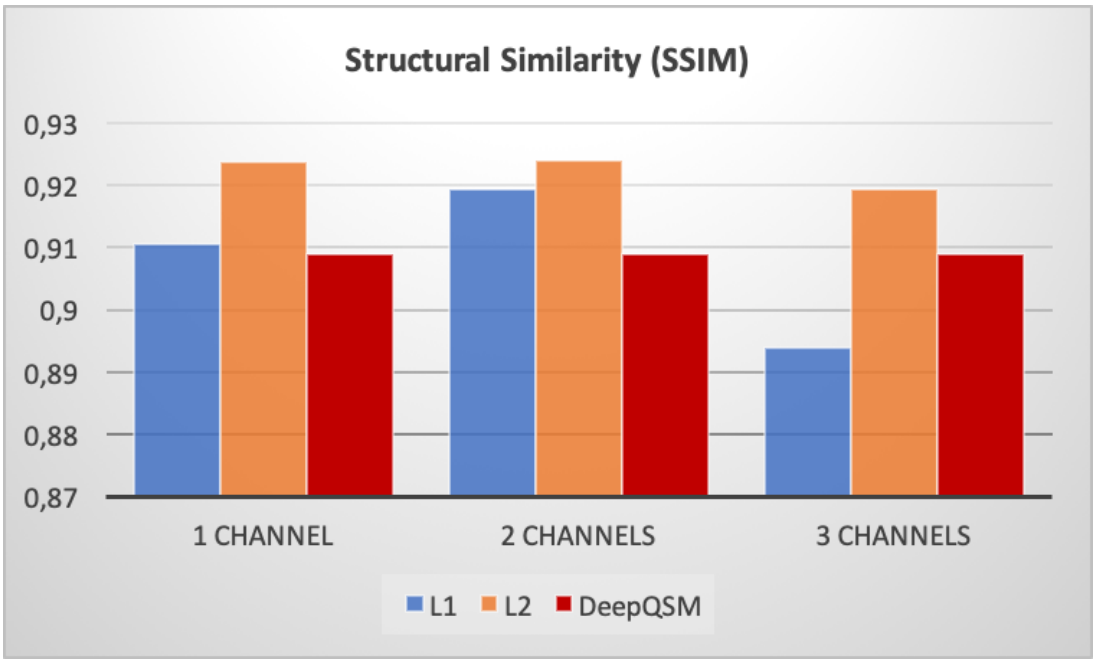


Figure 4.21: SSIM computed between on the predicted brain image and the lable used during the validation process. The blu bar shows the outcome the network using L1 as loss function, the orange bar shows L2 loss function based network and the red bar shows the MSE computed using DeepQSM from [38].

Afterwards, MSE has been calculated just within the corpus callosum, where high anisotropic effect is expected much more since white matter fibers are packed densely in this region. In figure 4.22 is shown the result of the segmentation.

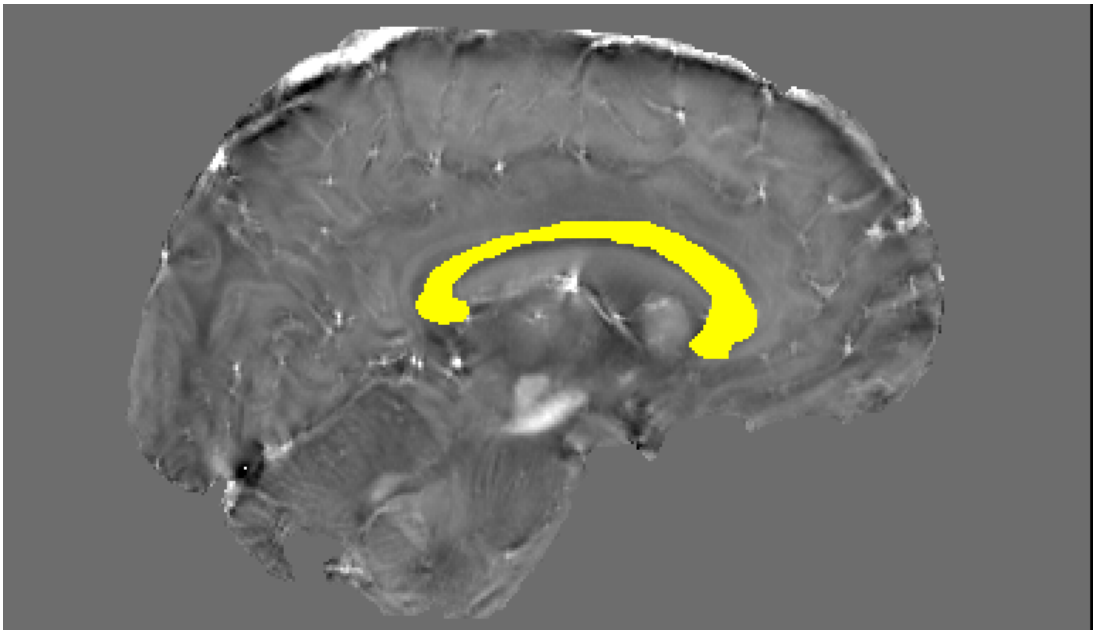


Figure 4.22: In yellow the corpus callosum, manually segmented using ITK-SNAPE.

Then MSE has been calculated in two different scenarios, where the input has been thresholded and where it has not. Figure 4.23 shows the outcome related to the prediction where the input had been thresholded, while figure 4.24 refers to the other type of input (in figure 4.3 are illustrated both type of inputs).

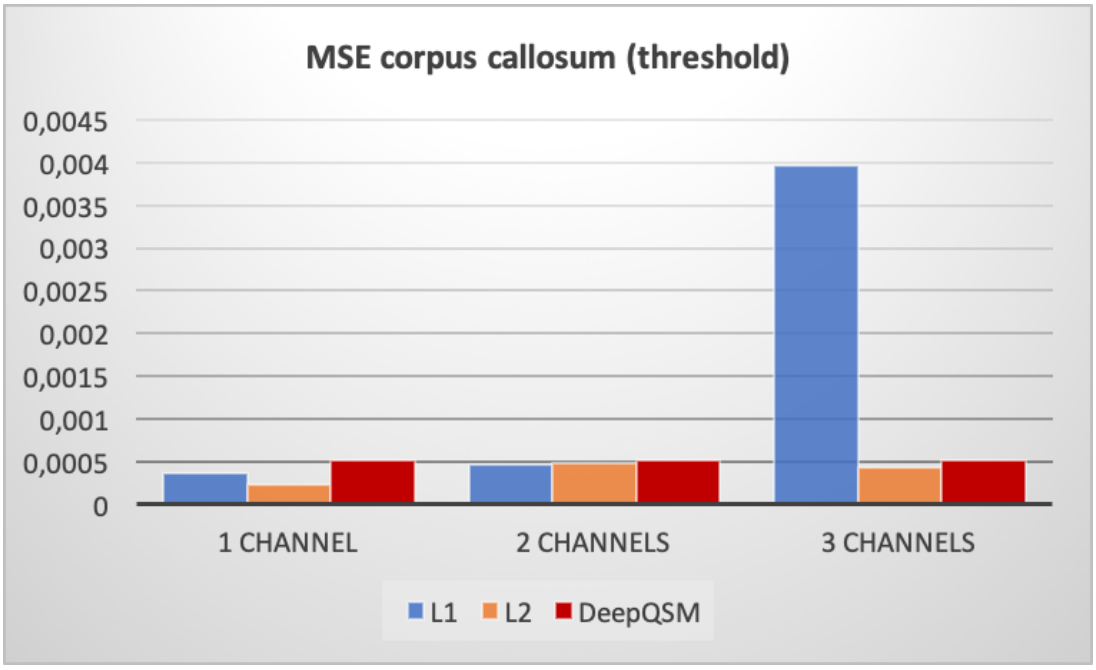


Figure 4.23: MSE calculated along the corpus callosum. The similarity refers to predictions when the input was thresholded. L1, L2 loss function have been compared and the red bar refers to the MSE computed when DeepQSM predicts the brain image.

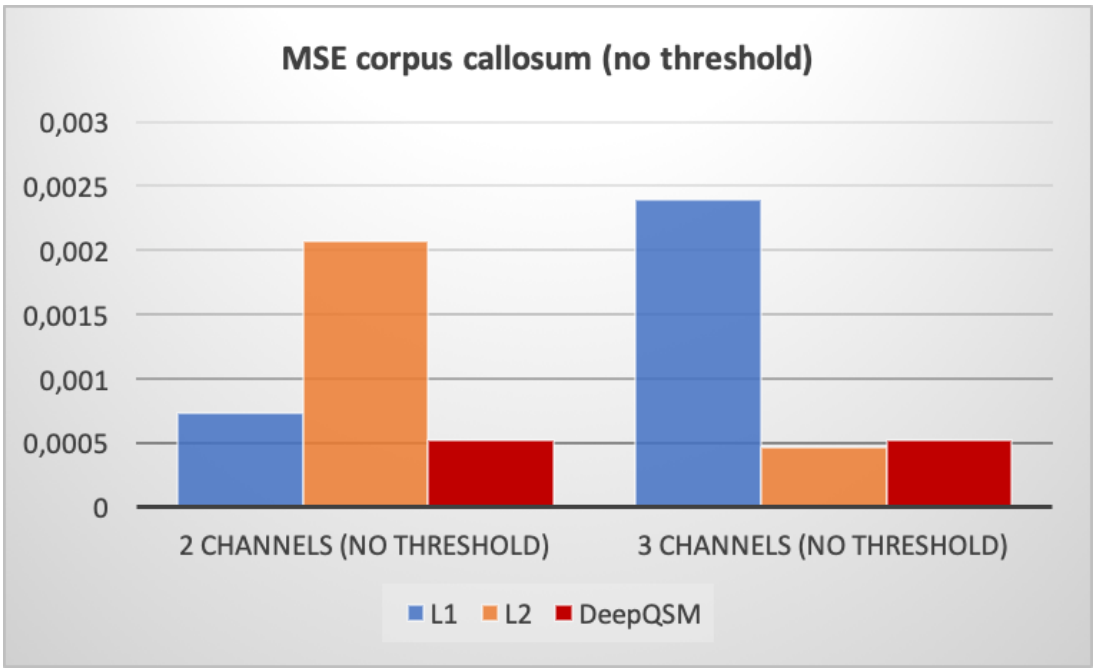


Figure 4.24: MSE calculated along the corpus callosum. The similarity refers to predictions when the input was not thresholded. L1, L2 loss function have been compared and the red bar refers to the MSE computed when DeepQSM predicts the brain image.

Chapter 5

Discussion

The model chosen in this project aimed to simulate white matter fibers (as longitudinal structures in the volume samples which are part of the training set). We were expecting the network to learn the anisotropic effect wherever it occurs (if FA is close to 1) and how this perturbation occurs (following a $\sin^2$ relationship). Unfortunately the validation has not been successful as we were expecting. One of the reasons why might be the low subject availability to validate the network (in this work just 1 subject has been chosen). This lack of data is correlated to the time-consuming process required by acquiring the two types of images (COSMOS and DWI) within the same session.

However, after the training has been performed, figures 4.4, 4.5, 4.8, 4.9, 4.12, 4.13 show how the loss function decrease in every training session. Unfortunately to evaluate the network some predictions had to be performed. Figures 4.6, 4.7, 4.10, 4.11, 4.14, 4.15 show the prediction of a brain image where the input has been thresholded (see figure 4.3). You can clearly see an unexpected effect in the voxels that had been thresholded in FA and orientation maps. Figure 5.1 shows the artifact where the FA has been thresholded.

For this reason we chose to use an input that were not thresholded, hoping that this artifact disappeared.

The first thing we want to highlight is whether L1 or L2 is the best loss function to use to solve a dipole-inversion problem. Figures 4.20 and 4.21 show MSE and SSIM calculated along the volume between the prediction and the label in order to evaluate which loss function was better, comparing them with the MSE and SSIM calculated

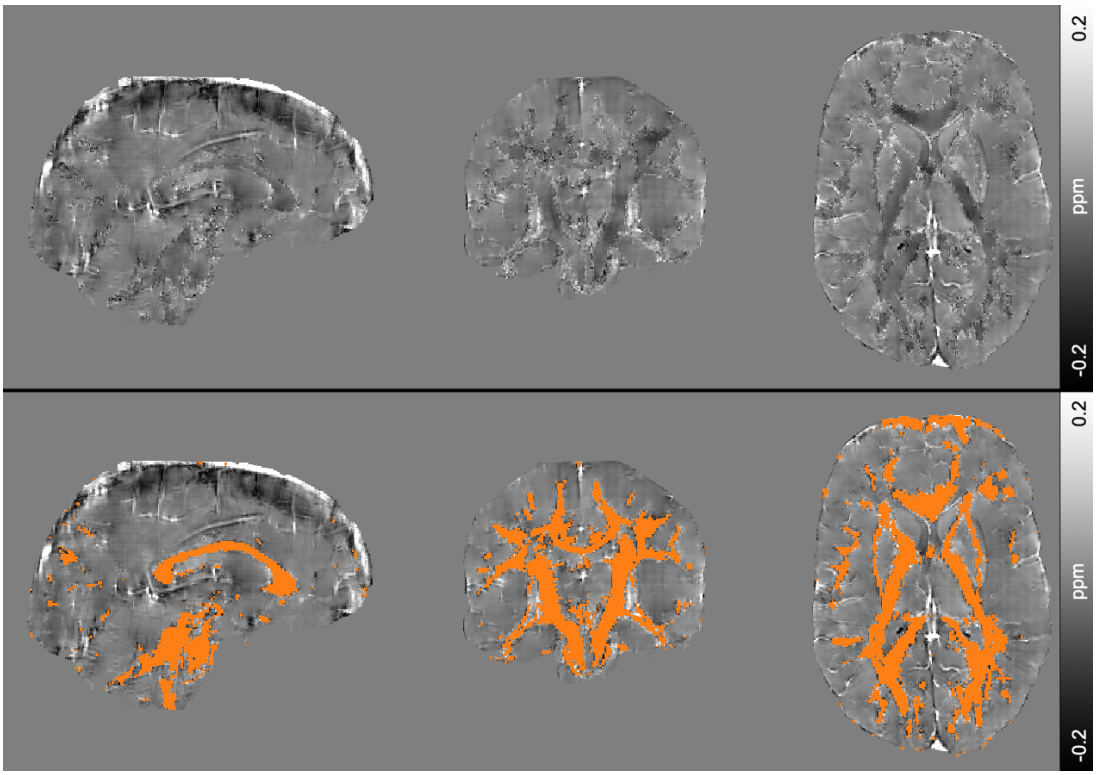


Figure 5.1: Prediction from 2 channel input architecture using L2 (top) and the same image when FA map (orange) is overlayed. It is clear the artifact coming from the thresholded input image.

using DeepQSM from [38]. Both similarity measures illustrate that L2 has provided a less prediction error and a better reconstruction. L2 is more robust than L1 as we were expecting. However, despite MSE shows a less error in 1 channel using L2 loss function, the equivalent prediction is not that good as the prediction from DeepQSM, which shows a better anatomical detailed representation. This is addicted to the fact that both MSE and SSIM are very sensitive to some factors such as shape borders presented in the volume and data range. So even though the shapes are maintained confident between the prediction and the label, the values are not that good, MSE would show a good similarity measurement.

Afterwards, MSE has been computed within the voxel delimiting the corpus callosum. Inside this brain region a high anisotropic effect is expected and no borders between different brain tissue are expected. So MSE should reveal a more reliable similarity measure. In this case, MSE has been computed between the predictions coming from the different channel architecture, using a threshold and a non-threshold input. Figure 4.23

and 4.24 compare such predictions. From these graphs it seems that just 1 channel architecture shows better results having a less MSE using both L1 and L2 loss function.

Finally, we can conclude that our proposal deep learning neural network seemed not to provide better results than before, except if we consider the comparison performed just within the corpus callosum. Where the anisotropic effects is expected, our neural network (1 channel, L2 loss function was the one with the least MSE) showed a better prediction than DeepQSM. This means that the training set model suggested in this work might bring further and useful information to the network to learn. However, the validation should be performed using more than one subject as in this work, so more reliable results could be studied if we validated a bigger dataset with brain images.

Chapter 6

Conclusion

Machine learning is part art and part science. When you look at machine learning algorithms, there is no one solution or one approach that fits all. Some problems are specific and require a unique approach. Some other problems are more complicated and need a trial & error approach. Supervised learning, classification and regression are very open.

In this work machine learning has tried to solve an inverse problem: one challenging aspect is handling the multivalued branches of the direct problem and discarding unphysical solutions. Exploring different trainings has been fundamental to understand which model can better fit the problem we want to achieve. As in every field, when machine learning has been chosen to invert the problem to the solution, a good and robust model has to be built. Choosing the approach is a critical step to achieve a promising performance. There is no deterministic method that might help choosing the right approach. Different approaches helped us to have a better idea of some future work that could be done based on this thesis.

Results from predictions were not that good as we were expecting - images did not have a good anatomical resolution as DeepQSM's prediction. Anyway, MSE performed within the corpus callosum demonstrates how the network might be able to reconstruct the anisotropic effect which was one of the challenges of this thesis. Maybe a more complicated model should be defined, considering fiber overlapping or different strategies. What we were not expecting from the predictions was an improvement difference between DeepQSM prediction and 1-channel input from our architecture: the only thing that was changing between the two scenarios was adding longitudinal structures to the volume

samples used during the training. We were not expecting so much difference from the anatomical details point of view, despite 1 channel input was the architecture that showed the less MSE compared to the others.

However, convolutional neural networks are known to be sensitive to the input data range and this might be one reason why we could not achieve high performance. Different data scaling might be proposed for the input during the training and the validation and, certainly, validating the network using more than one subject. This will give a more reliable way of measuring the accuracy of the network since one subject might not be enough to validate a neural network.

Bibliography

[1] Diffusion mri. http://markus-schirmer.com/old/mri_diffusion.html. Accessed: 2018-09-20.

[2] Diffusion tensor. http://mriquestions.com/diffusion-tensor.html. Accessed: 2018-09-25.

[3] Talin Babikian, M Catherin Freier, Karen A Tong, Joshua P Nickerson, Christopher J Wall, Barbara A Holshouser, Todd Burley, Matt L Riggs, and Stephen Ashwal. Susceptibility weighted imaging: neuropsychologic outcome and pediatric head injury. *Pediatric neurology*, 33(3):184–194, 2005.

[4] Vinit Baliyan, Chandan J Das, Raju Sharma, and Arun Kumar Gupta. Diffusion weighted imaging: Technique and applications. *World journal of radiology*, 8(9):785, 2016.

[5] Congbo Cai, Xi Chen, Weijun Liu, Shuhui Cai, Delu Zeng, and Xinghao Ding. Rapid reconstruction of quantitative susceptibility mapping via improved 0 norm approximation. *Computers in biology and medicine*, 79:59–67, 2016.

[6] Yu-Chung N Cheng, Jaladhar Neelavalli, and E Mark Haacke. Limitations of calculating field distributions and magnetic susceptibilities in mri using a fourier based method. *Physics in Medicine & Biology*, 54(5):1169, 2009.

[7] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.

[8] Andreas Deistung, Ferdinand Schweser, and Jürgen R Reichenbach. Overview of quantitative susceptibility mapping. *NMR in Biomedicine*, 30(4):e3569, 2017.

[9] Andreas Deistung, Ferdinand Schweser, Benedikt Wiestler, Mario Abello, Matthias Roethke, Felix Sahm, Wolfgang Wick, Armin Michael Nagel, Sabine Heiland, Heinz-Peter Schlemmer, et al. Quantitative susceptibility mapping differentiates between blood depositions and calcifications in patients with glioblastoma. *PloS one*, 8(3):e57924, 2013.

[10] Jeff Duyn. Mr susceptibility imaging. *Journal of magnetic resonance*, 229:198–207, 2013.

[11] Jeff H Duyn and John Schenck. Contributions to magnetic susceptibility of brain tissue. *NMR in Biomedicine*, 30(4):e3546, 2017.

[12] E Mark Haacke and Jürgen R Reichenbach. *Susceptibility weighted imaging in MRI: basic concepts and clinical applications*. John Wiley & Sons, 2014.

[13] Bo Hallgren and Patrick Sourander. The non-haemin iron in the cerebral cortex in alzheimer's disease. *Journal of neurochemistry*, 5(4):307–310, 1960.

[14] Naying He, Huawei Ling, Bei Ding, Juan Huang, Yong Zhang, Zhongping Zhang, Chunlei Liu, Kemin Chen, and Fuhua Yan. Region-specific disturbed iron distribution in early idiopathic p arkinson's disease measured by quantitative susceptibility mapping. *Human brain mapping*, 36(11):4407–4420, 2015.

[15] Xiang He and Dmitriy A Yablonskiy. Biophysical mechanisms of phase contrast in gradient echo mri. *Proceedings of the National Academy of Sciences*, 106(32):13558–13563, 2009.

[16] Yen-Lin Huang, Ying-Sheng Kuo, Ying-Chi Tseng, David Yen-Ting Chen, Wen-Ta Chiu, and Chi-Jen Chen. Susceptibility-weighted mri in mild traumatic brain injury. *Neurology*, 84(6):580–585, 2015.

[17] John David Jackson. *Classical electrodynamics*. John Wiley & Sons, 2007.

[18] Varsha Jain, Osheiza Abdulmalik, Kathleen Joy Propert, and Felix W Wehrli. Investigating the magnetic susceptibility properties of fresh human blood for noninvasive

oxygen saturation quantification. *Magnetic resonance in medicine*, 68(3):863–867, 2012.

[19] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[20] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[21] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[22] Denis Le Bihan, Eric Breton, Denis Lallemand, Philippe Grenier, Emmanuel Cabanis, and Maurice Laval-Jeantet. Mr imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2):401–407, 1986.

[23] Steven M LeVine. Iron deposits in multiple sclerosis and alzheimer's disease brains. *Brain research*, 760(1-2):298–303, 1997.

[24] Wei Li, Alexandru V Avram, Bing Wu, Xue Xiao, and Chunlei Liu. Integrated laplacian-based phase unwrapping and background phase removal for quantitative susceptibility mapping. *NMR in Biomedicine*, 27(2):219–227, 2014.

[25] Wei Li and Chunlei Liu. Comparison of magnetic susceptibility tensor and diffusion tensor of the brain. *Journal of neuroscience and neuroengineering*, 2(5):431–440, 2013.

[26] Wei Li, Chunlei Liu, Timothy Q Duong, Peter CM van Zijl, and Xu Li. Susceptibility tensor imaging (sti) of the brain. *NMR in Biomedicine*, 30(4):e3540, 2017.

[27] Wei Li, Bing Wu, Alexandru V Avram, and Chunlei Liu. Magnetic susceptibility anisotropy of human brain in vivo and its molecular underpinnings. *Neuroimage*, 59(3):2088–2097, 2012.

[28] Chunlei Liu. Susceptibility tensor imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 63(6):1471–1477, 2010.

[29] Chunlei Liu, Nicole Murphy, and Wei Li. Probing white-matter microstructure with higher-order diffusion tensors and susceptibility tensor mri. *Frontiers in integrative neuroscience*, 7:11, 2013.

[30] Tian Liu, Pascal Spincemaille, Ludovic De Rochefort, Bryan Kressler, and Yi Wang. Calculation of susceptibility through multiple orientation sampling (cosmos): a method for conditioning the inverse problem from measured magnetic field map to susceptibility source image in mri. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 61(1):196–204, 2009.

[31] Hendrik Antoon Lorentz. *The Theory of Electrons and Its Applications to the Phenomena of Light and Radiant Heart: A Course of Lectures Delivered in Columbia University, New York in March and April, 1906*, volume 29. Teubner, 1916.

[32] J Luo, X He, and DA Yablonskiy. Magnetic susceptibility induced white matter mr signal frequency shiftsexperimental comparison between lorentzian sphere and generalized lorentzian approaches. *Magnetic resonance in medicine*, 71(3):1251–1263, 2014.

[33] JP Marques and R Bowtell. Application of a fourier-based method for rapid calculation of field inhomogeneity due to spatial variation of magnetic susceptibility. *Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering: An Educational Journal*, 25(1):65–78, 2005.

[34] Masakazu Matsugu, Katsuhiko Mori, Yusuke Mitari, and Yuji Kaneda. Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks*, 16(5-6):555–559, 2003.

[35] Amanda Ching Lih Ng. *Techniques for the processing and analysis of magnetic resonance imaging phase data*. PhD thesis, 2013.

[36] Boris B Parfenenkov and Maksim A Panachev. Comparison of some image quality approaches. In *AIST (Supplement)*, pages 48–53, 2014.

[37] Daniel M Patterson, Anwar R Padhani, and David J Collins. Technology insight:

water diffusion mria potential new biomarker of response to cancer therapy. *Nature Reviews Clinical Oncology*, 5(4):220, 2008.

[38] Kasper Gade Botker Rasmussen, Mads Janus Kristensen, Rasmus Guldhammer Blendal, Lasse Riis Ostergaard, Maciej Plocharski, Kieran O'Brien, Christian Langkammer, Andrew Janke, Markus Barth, and Steffen Bollmann. Deepqsm - using deep learning to solve the dipole inversion for mri susceptibility mapping. *bioRxiv*, 2018.

[39] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[40] John F Schenck. The role of magnetic susceptibility in magnetic resonance imaging: Mri magnetic compatibility of the first and second kinds. *Medical physics*, 23(6):815–850, 1996.

[41] Ferdinand Schweser, Andreas Deistung, and Jürgen R Reichenbach. Foundations of mri phase imaging and processing for quantitative susceptibility mapping (qsm). *Zeitschrift für medizinische Physik*, 26(1):6–34, 2016.

[42] Alexander L Sukstanskii and Dmitriy A Yablonskiy. On the role of neuronal magnetic susceptibility and structure symmetry on gradient echo mr signal formation. *Magnetic resonance in medicine*, 71(1):345–353, 2014.

[43] Samuel Wharton and Richard Bowtell. Effects of white matter microstructure on phase and susceptibility maps. *Magnetic resonance in medicine*, 73(3):1258–1269, 2015.

[44] Ivo Wolf, Marcus Vetter, Ingmar Wegner, Thomas Böttger, Marco Nolden, Max Schöbinger, Mark Hastenteufel, Tobias Kunert, and Hans-Peter Meinzer. The medical imaging interaction toolkit. *Medical image analysis*, 9(6):594–604, 2005.

[45] Luke Xie, Russell Dibb, Gary P Cofer, Wei Li, Peter J Nicholls, G Allan Johnson, and Chunlei Liu. Susceptibility tensor imaging of the kidney and its microstructural underpinnings. *Magnetic resonance in medicine*, 73(3):1270–1281, 2015.

[46] Dmitriy A Yablonskiy, Xiang He, Jie Luo, and Alexander L Sukstanskii. Lorentz sphere versus generalized lorentzian approach: What would lorentz say about it? *Magnetic resonance in medicine*, 72(1):4–7, 2014.

[47] Dmitriy A Yablonskiy and Alexander L Sukstanskii. Lorentzian effects in magnetic susceptibility mapping of anisotropic biological tissues. *Journal of Magnetic Resonance*, 292:129–136, 2018.

[48] Jaeyeon Yoon, Enhao Gong, Itthi Chatnuntawech, Berkin Bilgic, Jingu Lee, Woojin Jung, Jingyu Ko, Hosan Jung, Kawin Setsompop, Greg Zaharchuk, et al. Quantitative susceptibility mapping using deep neural network: Qsmnet. *NeuroImage*, 179:199–206, 2018.

Acknowledgements

After six months in Australia, the day has come; writing my acknowledgements. This makes my thesis truly complete. It has been a period of incredible development, not only from a scientific point of view but also personally. Masters thesis represents the end of a very important phase in every university students career and the beginning of what hopefully might be a propitious and brave journey into the "adult world".

Special thanks to my professor Gabriella Balestra from Politecnico of Turin for supporting me in the choice of this project. My deepest gratitude goes to Leigh Johnston and Steffen Bollmann from University of Melbourne and the University of Queensland without whom this work would have never existed. Thank you for giving me the possibility to expand my knowledge in the fascinating field of neuroscience and to develop this master thesis project. Thank you to the post-doc Warda Najam. Her constructive comments and wise suggestions meant I learnt a lot from her and her guidance helped me improve my skills and determination in the most challenging times. I am eternally grateful to the countless times she helped me. The successful completion of this thesis is entirely due to their innovative ideas, critical analysis, excellent advice and detailed discussions during this time.

To my family, thank you for encouraging me in all of my pursuits and inspiring me to follow my dreams. I am especially grateful to my parents, who supported me emotionally and financially. I always knew that you believed in me and wanted the best for me.

I would like to thank all my faithful friends in Turin that assisted and encouraged me. I'm also deeply indebted to Leonardo for supporting me from overseas every time I needed encouragement to do my best. He taught me to always look ahead and keep fighting for the pursuit of my dreams. Special thanks goes to my old friends Daniele, Elia, Francesco, Luca, Matteo e Paolo which whom I spent most of the good times in Turin, and also to my flatmates whom, even if I've known for a short time, they always made me feel at home. Alice, Tommaso (friend of adventures and misadventures) and Manuela; thank you for being my second home in Turin.

Particularly helpful to me during this time was my dear friend Federica, whom I hold in very high regard. Thank you for listening to me when I was experiencing challenges

along the way. Thank you also to Sofia Pina, my oldest friend. I have always considered you a true friend. I know that I can be always myself around you and for that I am grateful. Special thanks to Antonella because she understands me, I know I can count on her every time I need. Many thanks to Benedetta because after all these years our friendship is still stronger than ever. To all my friends and colleagues with whom I spent countless days in Turin between exams stress and fun. To Alessia, Alma, Carla, Flavio, Francesca De Meis, Gerardina, Mario, Marta, Mattia, Michele, Paolo Borlengo e Simone.

Lastly, but not less importantly, to my family in Australia. Thank you for loving and caring for me unconditionally during my stay in Australia. You really made me feel part of your family and thank you for making my time in Melbourne pleasant and enjoyable.

Giampaolo Tuccio