NUMERICAL BODY MODEL INFERENCE

FOR INDIVIDUALIZED RF EXPOSURE PREDICTION

IN NEUROIMAGING AT 7T MRI

TU Delft Delft University of Technology

Challenge the future

NUMERICAL BODY MODEL INFERENCE FOR INDIVIDUALIZED RF EXPOSURE PREDICTION IN NEUROIMAGING AT 7T MRI

by

Prernna Bhatnagar

in partial fulfillment of the requirements for the degree of

Master of Science
in Electrical Engineering

at the Delft University of Technology,
to be defended publicly on Thursday October 29, 2020 at 9:00 AM.

Student number:	4772881
Supervisor:	dr. ir. R.F. Remis, CAS, TU Delft
Daily Supervisor:	dr. ir. W.M. Brink, C.J. Gorter Center, LUMC
Thesis committee:	dr. ir. R.F. Remis, CAS, TU Delft
	dr. ir. M Staring TU Delft and LUMC, LKEB
	dr. ir. W.M. Brink, C.J. Gorter Center, LUMC
	dr. ir. Sahar Yousefi, C.J. Gorter Center, LUMC

This thesis is confidential and cannot be made public until October 29, 2020.

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft Delft University of Technology

LUMC Leids Universitair Medisch Centrum

ACKNOWLEDGEMENT

"I am one of those who think like Nobel, that humanity will draw more good than evil from new discoveries."- Marie Skłodowska Curie

I would like to extend my thanks to my daily supervisor from LUMC(Leids Universtair Medisch Centrum) dr.ir. W.M. Brink for his guidance, support and relentless patience even when things were not going right and for familiarzing me with scanning at MRI 7T scanner. I would like to thank dr.ir. Sahar Yousefi from LUMC for being patient and guiding me through the deep learning aspect of the project. A special thanks to dr. ir. Rob Remis, my supervisor at TU Delft for making the entire process smooth and for providing your valuable insight and feedback. I would also like to extend thanks to dr. ir. Marius Staring for agreeing to be a part of my thesis committee.

I have heartfelt gratitude towards my parents and brother who always encouraged me to follow my dreams and never give up in times of trouble, they have been a constant support for me. I would also like to thank my friends in Delft and in India for being there throughout.

Last but not the least I would like to thank my dog, Benz and all the numerous dogs I met on walks!

As challenging as these times maybe, I hope they would be remembered more as an epitome of scientific breakthroughs and discoveries, human resilience and creativity and less for loss, tragedy and trauma.

Prernna Bhatnagar
Delft, October, 2020

ABSTRACT

Magnetic Resonance Imaging is a popular modality for brain imaging in present times. The quality of the images depends on the strength of the magnetic field. An MRI scanner with a magnetic field strength of 3 Tesla(T) is pre-dominantly used for clinical purposes. However, with the advancement in technology, and the need to image finer image finer structures, we are gradually shifting to higher magnetic field strengths like 7T and above. One of the major bottlenecks in these systems is the bias induced in the images due to field inhomogeneities of higher field strengths. Yet another drawback of these systems is the increase in power dissipation in the tissues. This is measured by a quantity called Specific Absorption Rate(SAR). The goal of this thesis is to accurately predict SAR values for various volunteers as they may differ from subject to subject. Many homogeneous models have been created earlier to estimate the value of SAR, however, these estimates are often over-conservative and safe which compromises with image quality. A personalised numerical body model is created for all volunteers using relevant information derived from simulations performed on generalized models. Various tissues have to be segmented to create these 3D numerical body models. However, there has to be a trade-off between ease of segmentation and SAR accuracy. Keeping that in mind, it was found that the optimum number of tissues to get reliable SAR estimates is six. A deep learning method was then used for segmentation. A numerical body model was derived for all the volunteers using the deep learning segmentation. An adapted ForkNet, which is similar to U-net in architecture is used to segment these images. The SAR values derived from the predicted numerical body model and the original body model are similar, hence speeding up the process for SAR prediction. However, there are certain limitations of the thesis that can be addressed in the future. Inadequate data remains a major bottleneck for the project, increasing the data should result in improved segmentation, this can be addressed by acquiring more data. Another major drawback of the thesis is the segmentation accuracy, the ground truth segmentation is performed to the best of our knowledge, however, some errors are still present in the ground truth segmentation. The next steps for this project would be to acquire more data, train the data on multiple input sequences, use a 3D network and using localizer images that are acquired at the start of the MRI scan. Nonetheless, the principles established in this thesis confirm that a deep learning approach can be used to create numerical body models for SAR estimates. It also establishes the fact that these SAR estimates are comparable to the SAR estimates generated from the ground truth numerical body models.

Contents

1

INTRODUCTION

This chapter introduces the main objective and motivation behind the thesis. Section 1.1 highlights basic MRI physics. This is followed by the motivation behind the thesis in Section 1.2. Section 1.3 presents the main contribution of the thesis and Section 1.4 highlights the state-of-the-art and prior research in the field. Lastly, Section 1.5 discusses the organization of the thesis.

1.1. MAGNETIC RESONANCE IMAGING

Magnetic Resonance Imaging (MRI) is a non-invasive, non-ionizing imaging modality. It is based on the principle of rotating spins in the human body and their behaviour when they experience a magnetic field. It is used to detect tumors in the brain, in musco-skeletal injuries and to image brain tissues.[1],[2]

Since its inception, it has been widely used to diagnose brain-related disorders due to its ability to image soft tissues with good spatial resolution and contrast. This aids in distinguishing between healthy tissue and a tumor(which can be malignant or benign). Since MRI scanners can be used for diagnosing not just cancer but many other diseases in the human body, it has increased the care that doctors can provide to their patients.

The human body consists of 70% water and therefore contains an abundance of protons that are fundamental to generating the MRI signal. The spins(hydrogen protons) inside the human body act as tiny magnets. They are arranged randomly before the application of an external magnetic field. When the static magnetic field is applied($\mathbf{B}_0$), the spins align themselves to the external magnetic field and can be perturbed using a radio frequency(RF) pulse. When the RF pulse is applied the spins start precessing with the resonant frequency of the static magnetic field, also known as Larmor frequency. When an RF pulse is turned off, these spins return to their equilibrium state. During this process, the spins induce a voltage in nearby RF coils which can be used to create an image.

Two types of magnetic fields influence the spins. The RF magnetic field is known as the B_1 field and the static magnetic field is known as B_0. The $\mathbf{B}_1$ field can be separated into oscillating linearly polarized components B_{1x} and B_{1y} , they can also be expressed as circularly polarized components rotating clockwise(B_1^+) and counter-clockwise(B_1^-) within the transverse plane. The relationship between the rotating components, B_1^+ (appearing to rotate in a counter-clockwise direction when looking 'down' at it from a location in the +z-direction) and B_1^- (appearing to rotate in a clockwise direction when looking 'down' at it from a location in the +z-direction), and the linearly polarized components B_{1x} and B_{1y} can be expressed as:

$$B_1^+ = \frac{(B_{1x} + iB_{1y})}{2} \tag{1.1}$$

$$B_1^- = \frac{(B_{1x} - iB_{1y})*}{2} \tag{1.2}$$

One of the two rotating components will rotate in the same direction as nuclear spin, and will thus, match the Larmor frequency. This appears as a static field in the frame of reference of the spin, rotating about z axis

and therefore causes spin mutation. This component is referred to as the B_1^+ field.

Th RF field also induces current in certain parts of the body. Tissues that have high electrical conducitvity are more susceptible to these currents. According to Ohm's Law:

$$J = \sigma E \tag{1.3}$$

Where J is the current density, σ is the conductivity and E is the electric field. This current leads to power deposition in tissues which is converted into heat. Power gets dissipated in human tissue as thermal energy. Therefore, excessive RF power deposition in an MRI experiment is unsafe since it will cause a temperature increase in the body of the subject. Widely heterogeneous properties of tissues make it harder for predicting the exact rise in temperature. However, this can be established by the relationship between thermal power and Specific Absorption Rate(SAR) is defined as the power deposited per unit tissue mass (units of Watts per kilogram).

$$SAR = \frac{\sigma |\mathbf{E}|^2}{\rho} \tag{1.4}$$

Where, σ denotes the electrical conductivity, $\mathbf{E}$ denotes the electric field strength and ρ denotes density. Hence, SAR is dependent on the electrical field, the conductivity of the tissue, the density of the tissue and also the relative electrical permitivitty of the tissue. It has become the standard metric for describing RF safety in MRI. SAR is characterized as a global and a local phenomenon. Delivering power to a subject will cause the average temperature to increase which is limited through the whole body. However, given that SAR arises from the local electric field, some tissue volumes inside the body will have higher SAR than others, leading to non homogeneous heating patterns. To limit the local RF exposure, local SAR limits are defined in terms of 10g-averaged SAR.

1.2. MOTIVATION

Most MRI scanners used for clinical applications operate at a strength of 1.5T and 3T. However, they suffer from low signal-to-noise ratio (SNR) which limits the attainable spatial resolution. High SNR guarantees high spatial resolution and finer structures of the human body can be measured.
Since the SNR increases with increasing B_0 field, MRI 7T scanners would be used for obtaining a higher spatial resolution that are capable of imaging small structures in comparison with the 3T scanner. A specific example in neuroimaging for 7T is in the cerebral cortex for the detection of changes in cortical structure, like the visualization of cortical microinfarcts and cortical plaques in Multiple Sclerosis.[3]. Vascular imaging is another highly promising application for 7T.[3]
In ultra-high field(UHF) MRI (>3T), the wavelength of the RF field becomes comparable to the width of the human body. This causes inhomogeneous excitation patterns, with varying flip angles (and thus signal intensity/contrast) over the field of view. To solve this problem, we can use several smaller transmit coils in parallel(parallel transmit or pTx). Each coil has a different transmit sensitivity which provides better coverage. Phase and amplitude of the array elements can be adjusted in many ways to provide a wide range of B_1^+ field patterns and mitigate the excitation non-uniformity. However, this comes at the cost of increased SAR. Furthermore, pTx arrays typically consist of smaller elements placed closer to the body, which may also cause SAR to accumulate locally near the elements. Given that at 7T, there is already an increased chance of SAR related tissue heating due to the higher RF frequency, it is imperative that pTx SAR should be accurately assessed for pTx to be considered safe and useful for improving image quality. [4]

RF power deposition, typically quantified by the specific absorption rate (SAR) in Watts per Kilogram(W/Kg), is a concern at all field strengths but becomes more limiting at 7T, necessitating longer scan times or sacrificing image quality.

To solve this problem, many numerical models have been widely used in the past to estimate SAR values. These values are calculated by considering the worst case scenario so these values are safe, but overestimated, hence the quality of the images is sub optimal. SAR values depend on the body composition, geometery and the position of the person inside the scanner. Since, the body models would not take these factors into consideration, they lead to SAR values that may not be optimal for a specific patient.

Systems are now constantly pushing for more SNR by using higher field strengths of 9.4T, 11T and 14T and over estimation of SAR will also remain a bottleneck in these systems. We can only get close to complete ex-

ploitation if SAR limits are based on accurate predictions, rather than on conservative worst case estimations.

1.3. RESEARCH QUESTIONS

The main **objective** of this thesis is to create a personalised numerical body model to estimate SAR using deep learning methods for images which are acquired at the 7T scanner. The sub goals of the thesis are:

1. **Multi-contrast Ground Truth**: In this thesis a multi-contrast data acquisition is proposed to make a reliable ground truth body model. The high-contrast body model ensures that all the anatomical details are incorporated in the ground truth data.

2. **Segmented Tissues**: This thesis proposes to segment a total of six tissues using the multi-contrast images to create accurate body models

3. **Deep Learning**: The thesis proposes a novel modification of U-net for automatic segmentation of tissues.

1.4. STATE-OF-THE-ART

1.4.1. BODY MODELS AND SAR

It is evident from 1.2 that SAR is overestimated in present MRI scanners. This unrealistic safety standard for MRI scanners remains a major bottleneck for the quality of images, especially at higher field strengths. There is a need for patient-specific RF exposure estimation to improve image quality.

In earlier studies,[5] a cylindrical system was considered as a body model on which simulations were performed. However, the human body is more complex than just a cylinder, having hundreds of tissues and very complex anatomy. To solve this problem Virtual Family Voxel based body models were created to estimate SAR using a human body model.[6]

Previous studies have addressed these problems by increasing the amount of MRI data available to create a body model.[7]This means more databases, more segmentations and more labeled data available to an MRI technician at the beginning of an MRI scan.[7]. This study also aimed at reducing the number of tissues to a desirable number without causing much change in local SAR hotspots. This study concluded that the results obtained are for 3T images and they cannot make similar predictions about 7T images. This is due to the change in wavelength of the radio-frequency due to higher magnetic field strengths.

In another study[8], the authors concluded that the segmentation of all tissues can be replaced by a three tissue model(Water(Muscle), Fat and Lungs). To check if all water-rich tissues in the body can be replaced with muscle, some simulations were performed. The simulations were carried out on the Visible Human Male and Female models using a) Original model(36 tissues in male and 33 tissues in female), b) Dielectric properties of bone were assigned to fat, and the properties of muscles assigned to all other tissues except for the lung(“muscle-fat-lung model”)and c)The fat segments are replaced by muscle as well(“muscle-lung model"). It was observed that SAR values for the full model and “muscle-fat-lung model” are very similar, but SAR values for the “muscle-lung model” increased, indicating that fat cannot be turned into muscle and is an important tissue for SAR estimations. This work re-enforced the need for development of body models for patient-specific and more accurate SAR simulations.

In another state-of-the-art study [9] a similar problem is addressed. Accurate estimation of SAR is difficult in generalized body models. Anatomical and physiological fluctuations in the actual patient can vary substantially from the generalized models. In this study, SAR simulation is performed only for the Dixon. Deep learning and computer-vision algorithms are used to synthesize water and fat fractions without acquiring the Dixon thus reducing the scan time. The authors used prior information segmentation algorithms and then using computer-vision techniques for segmentation. It was found that there was a noticeable difference between fat and water similarity metrics, showing that fat has greater variability than water. They also observed similar B_1 field patterns for all models of different subjects. It is also observed that local SAR distribution may vary within the head for different models. On the other hand, peak SAR values are very close. In

conclusion, it can be said that computer-vision and deep learning approaches can be used for segmentation which can be used for predicting patient-specific RF exposure.

1.4.2. MANUAL AND SEMI-AUTOMATIC SEGMENTATION

Most of the studies stated above have used 3T images for segmentation. 3T images are not greatly influenced by the B_1 field. However, images from the 7T scanner are prone to image inhomogeneities which remains a major bottleneck for segmentation at 7T. Although many studies discuss the segmentation of 3T images[10],[11], not much work has been done on 7T images. Some of the 3T segmentation algorithms suggest using a 7T image for segmentation and then using the segmentations from the 7T propagated to the space of 3T images. In this approach, the 7T images were first skull stripped using FSL's toolbox BET then segmented using the FSL(FMRIB, Oxford, UK) FAST Segmentation toolbox. There are some residual errors even after the segmentation which are then corrected manually using ITK-SNAP[12]. When the corresponding 3T images are created from the 7T ground truth, a random forest classifier algorithm is used to segment these tissues automatically.
In another approach [13] a 10 tissue model was segmented. Those tissues include grey matter, white matter, cerebrospinal fluid, eyes, fat, muscle, skin, blood, bone(cancellous) and bone(cortical). Most of these methods are manual or semi-automatic algorithms using region growing, k-means clustering algorithm, etc. Many open-source platforms like 3D Slicer[14], FSL(FMRIB, Oxford, UK), Freesurfer[15], ITK-SNAP[12] have been used to create these segmentation models.
A state-of-the-art algorithm to segment 7T MP2RAGE images is presented in [16]. The main goal of this study was to compare their algorithm against ground truth segmemtations from FSL(FMRIB, Oxford, UK), Freesurfer[15], and SPM12[17]. The main parameters to be considered were similarity metrics and computational time. The skin and skull are stripped from the brain using FSL'S(FMRIB, Oxford, UK) BET(Brain Extraction Toolbox). MP2RAGE INV2 image is used for this purpose because it has less background noise. The mask from the brain stripped image is then used on the MP2RAGE UNI image to create a brain mask that has good contrast between grey and white matter. The mask is further eroded to remove non brain structures. The final mask is used in FSL's(FMRIB, Oxford, UK) FAST segmentation toolbox to get white matter, grey matter and cerebrospinal fluid segmentation. The intensities of the whole brain were normalized after skull stripping. This is a simple and effective algorithm to segment the brain in 7T images. However, the segmentations at the borders of the images still poses a problem.
Since various manual and semi-automatic segmentation algorithms are cumbersome and time-consuming, automatic segmentation of these structures is needed to save time and effort. The next subsection discusses the evolution of computer vision and deep learning methods for segmentation.

1.4.3. DEEP LEARNING BASED METHODS

In the last few years, many anatomical models have been used in SAR studies for determining the amount of power dissipation in tissues.[5],[18],[7],[19]
Manual and semi-automatic segmentation techniques have been used to segment these tissues. However, these algorithms are labour-some and time consuming. Any potential error in segmentation can lead to in correct SAR values. Several methods have been proposed in the last decade to avoid the errors due to manual and semi-automatic segmentation. Deep learning techniques are now being used to segment the tissues. Some of these methods have also been used to estimate the dielectric and physical properties of tissues based an anatomical images. These techniques are now the state-of-the-art in pattern recognition and data labeling problems. The key factor for the success of these algorithms is the architecture that can facilitate better extraction of feature maps. In a recent study, [20], the authors use a deep learning architecture to extract the dielectric properties of the tissues that have to be segmented for SAR evaluation. The authors first segment the images to generate the annotated head model. Then tissues based electrical properties are assigned to each tissue. The deep learning architecture aims to estimate the conductivity and permittivity maps using normalized T1- and T2-weighted images.

In a similar work[21] the authors have developed a Deep Learning method using U-net[22] architecture to segment 13 tissues. U-net is a convolutional neural network(CNN) that is widely used in medical image segmentation. The ground truth segmentations are extracted in a similar way like [13]. The NAMIC 3T Multimodality Brain dataset is used for this purpose. They segment the following tissues: skin, muscle, fat, bone(cortical), bone(cancellous), dura, blood, cerebrospinal fluid, grey matter, white matter, cerebellum, vitreous humor and mucous. As already discussed above, semi-automatic segmentation algorithms have some

limitations, especially for non brain tissues. The proposed CNN architecture in this study connects a single input to multiple outputs(N=13) which is the number of different tissues. The input is a 2D MRI slice and the output is a a 2D label field, in which the labels are integer numbers between 1 and 13. The model was then trained for different slicing directions, axial, sagittal and coronal and then combined to create a 3D model. A comparison was drawn between the ground truth segmentation and the output of the CNN . The results show high dice coefficients for the ForkNet architecture as compared to the conventional U-net architecture. There are still limitations like training data limitations, computational limitations which may be further improved by tuning the optimization parameters.
Hence, it was observed using the above studies that accurate prediction of SAR values is possible using deep learning algorithms making it quicker and time consuming.

1.5. THESIS ORGANIZATION

The structure of the document is as follows: The thesis is divided into six chapters

Chapter 1 discusses the introduction, motivation for the thesis, research questions and state-of-the-art. **Chapter 2** discusses background of the thesis. This chapter discusses in detail the background behind numerical body models, segmentation and deep learning. **Chapter 3** discusses methodology. This chapter discuses in details the implementation of the thesis with respect to simulations of numerical body models, segmentation and creation of patient-specific numerical body models and deep learning. **Chapter 4** discusses the results obtained in the thesis. **Chapter 5** discusses the conclusions drawn from the thesis, limitations of the thesis and future work involved.

2

BACKGROUND

Section 2.1 of this chapter describes the numerical body models derived from MRI data. It discusses in detail how these models have progressed from simple models to more complex models being used presently. Section 2.2 describes the acquisition and image processing of MRI images. It discusses in detail how different MRI sequences have been obtained and the parameters used for acquiring them. Section 2.3 describes the Deep learning approach of segmentation of images and the related mathematics and concepts behind it

2.1. NUMERICAL BODY MODELS

2.1.1. INTRODUCTION TO NUMERICAL BODY MODELS

MRI systems are prone to heating due to the use of RF coils of high magnetic strength. This is an ever-growing problem especially in Ultra High Field(UHF) MRI scanners. The magnetic coil gradients tend to induce voltage and heating which creates an electric eddy current pathway in tissues and leads to local heating of the tissue. This heat dissipation parameter is calculated in terms of SAR which is explained in much more detail in chapter-1. To prevent underestimation of SAR values numerical body models are created so that the SAR levels produced during an MRI scan are safe. A numerical body model is a discretized form of the human body that is used for numerical simulations. Numerical simulations become an invaluable tool to predict SAR values. Information can be extracted from these simulations like the electric field produced inside the human body and the eddy currents generated. Maxwell's equations need to be solved to simulate the interaction between radio frequency(RF) and the human body. The equations can be discretized and solved using Finite Difference Time Domain (FDTD)[23]or Finite Integration Technique(FIT) which are both time domain solvers. Numerical models have long progressed from being geometrical models[24] to very detailed anatomical models existing today. These models have been derived from different modalities spanning over a diverse range of age, gender, body compositions as performed in the Visible Human Project.[25].

2.1.2. HISTORY OF NUMERICAL BODY MODELS

In the 1980's a cylinder was used as a body model for predicting SAR values[26], it was an underestimation as the composition of the human body is heterogeneous and much complex and cannot be represented as a cylinder. This led to the rise of creating voxelised body models like the Visible Human Male and Female[27]-[28], the Virtual Family [6], and the NORMAN and NAOMI models [29],[30]. The SAR estimates that are currently being used are obtained from human models like the ones mentioned above. However, these models are not sufficient and we need more numerical models to accurately estimate SAR values. Efforts to increase the diversity of these models have enabled SAR levels to be studied in a more general population-wide context, and derive safety margins that allow safe application of MRI in all subjects. These safety margins, however, have come to compromise the performance of UHF MRI. If we can accurately predict SAR values we would not have to compromise on the image quality of MRI images, especially the images acquired using UHF MRI scanners.

The Virtual Family Model[6] has been used in this thesis to study tissue reduction on patient-specific data. Chapter-1 describes the importance of tissue reduction for numerical simulations. In a 3T study conducted by [8] The Visible Human Model was reduced to a three tissue model namely lung, fat and muscle. However, there are tissues that have to be segmented as they influence SAR values. In this study the focus was on the

whole body. However, in the current thesis the focus would only be on the tissues of the head and not the whole body. The brain itself is made up of numerous cortical and sub-cortical structures that it is difficult to classify the complete brain as just one tissue. We need more diversity in terms of dielectric properties of various tissues rather than just fat and muscle. Since, SAR values are determined by the dielectric and physical properties of the tissues, we need more than just two tissues to correctly estimate these values. For instance, if we draw a comparison between grey matter and CSF, the electrical conductivity in CSF is higher than grey matter so it is more relevant for SAR estimates. Hence, it is important to find tissues that are important in terms of SAR.

In a similar study [31], the authors compared different types of tissue reductions by combining different kinds of tissues together. The study is performed using numerical models in a 7T birdcage coil. In the study mentioned above, three different models are considered, the first model consists of a two tissue model: low proton density tissues and high proton density tissues. The low-proton-density is the mixture of internal air, cartilage, cortical bone, bone marrow, and teeth. The high-proton-density equivalent is a mixture of the rest of the head tissues. An average value of the density of these tissues was considered to be the density parameter and a similar process followed for the dielectric properties. The two tissue model is a very simple model, especially as far as image processing is concerned.
In the second model(3 Tissue Model) fat was removed from high-proton-density equivalent tissues.
In the third model(4 Tissue Model), white matter, grey matter, cerebellum and nerve were removed from the remaining of the second model.

B_1^+ maps and 10g average and 1g average SAR were computed, it was observed that: 4 tissue model is better for Ella and Thelonius, while the three tissue model is better for the Duke and Hugo models. Except for the Duke model, the difference between the three and four tissue model is insignificant. The tissue clustering strategy in this study was based on the ease of image segmentation. This study proved that reducing the number of tissues to create numerical body models does not have a major effect on SAR values. To create body models, tissues have to be segmented from MRI images. The next sub-section explains the data acquistion for this thesis.

2.2. MRI ACQUISITION

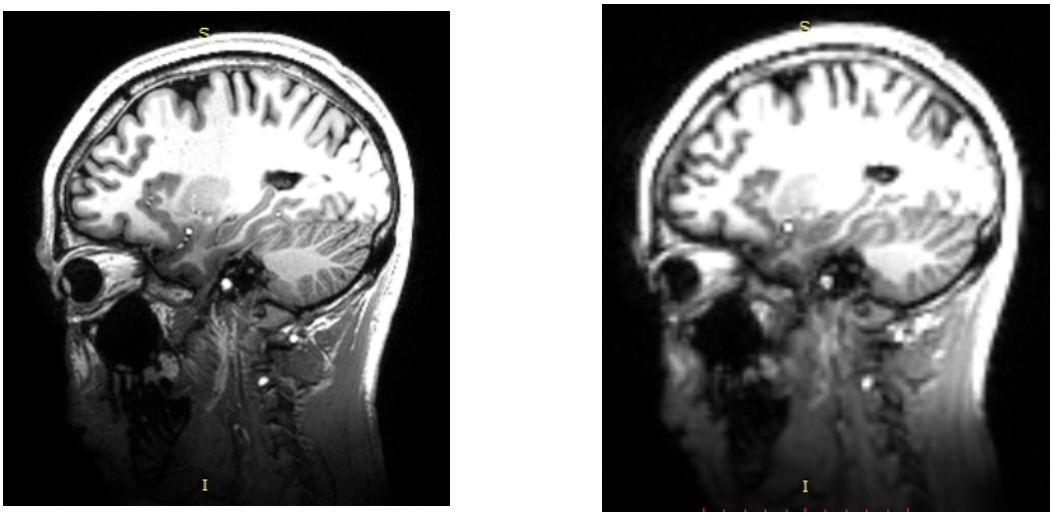
The images for this thesis were acquired on a Philips Achieva 7 Tesla(Philips Healthcare, Best, The Netherlands) scanner at the LUMC(Leids Uinversitair Medisch Centrum). The sequences were acquired using a quadrature transmit/receive birdcage coil (NM008A-7P-012; Nova Medical, Wilmington, MA) and a 32-channel receive array (NMSC075- 32-7P; Nova Medical, Wilmington, MA). The following sequences were obtained using the Philips Achieva 7 Tesla scanner. B_1^+ maps were also obtained to account for the inhomogeneity in 7T images.

S.No.	Sequence	Resolution	FOV
1	T1 weighted	1x1x1mm^3	256mmx256mmx192mm
2	T2 weighted	1x1x1mm^3	256mmx256mmx192mm
3	B1 Map	4x4x4xmm^3	256mmx256mmx192mm
4	Dixon	1x1x1mm^3	256mmx256mmx192mm
5	MP2RAGE	1x1x1mm^3	256mmx256mmx192mm
6	Proton-Density Weighted	1x1x1mm^3	256mmx256mmx192mm

Table 2.1: MRI Acquisition Protocol

One of the subgoals of this thesis is to obtain an accurate ground truth model, guided by the multi-contrast data.
T1 weighted Image:A tissue's T1 time reflects the amount of time its protons' spins realign with the main magnetic field(B_0). T1 weighted MRI images have short TE and TR times. Fat realigns with B_0 much faster than water, hence it appears brighter in a T1 weighted sequence. On the other hand, water has a much slower rate of re-alignment with the B_0 field, hence it appears darker. An isotropic resolution of 1mm was chosen to better capture the separation between tissues with otherwise similar MR contrast, such as grey matter and dura. This was found to be particularly important in the process of skull stripping.

T1 images were acquired at 1mm resolution for better segmentation quality.



(a) T1 image at 1mm

(b) T1 image at 2mm

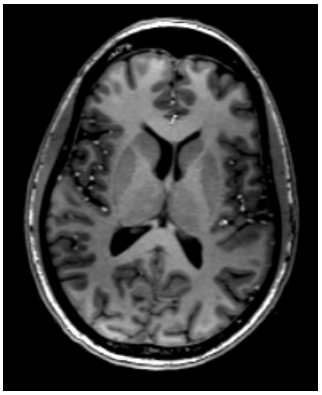
Figure 2.1: Resolution of T1 images

T2 weighted Image: A tissue's T2 time reflects the amount of time it takes for the protons to decay. A T2 weighted image requires long TR and TE times. In a T2 weighted image, water decays faster than fast, hence it appears bright while fat appears dark. Cerebrospinal Fluid(CSF) appears bright on a T2 weighted image, it provides a good contrast between brain tissues. This T2 weighted image is used to segment CSF in the brain. T2 images were acquired at 1mm resolution for better segmentation quality.

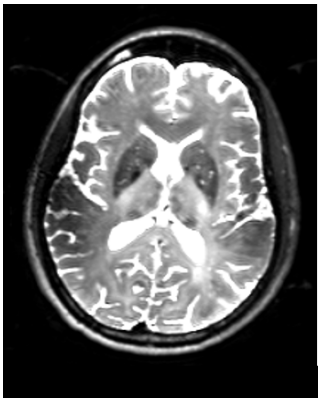
Dixon:The Dixon sequence is an MRI sequence that is generally used for fat suppression. The Dixon sequence exploits the fact that fat and water molecules precess at different rates. When they precess at different rates they keep shifting between being in-phase and out-of-phase, simultaneously acquiring both kinds of images. Dixon images were acquired at 1mm resolution for better segmentation quality.

Proton Density(PD) Weighted Image: A PD weighted image is an intermediate sequence sharing some features of both T1 and T2 images. Two variants of the PD weighted images are acquired, namely Turbo Field Echo(TFE) and Fast Field Echo(FFE). Both of these names are commercial names for a gradient echo sequence.

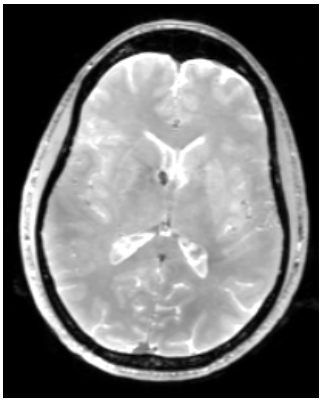
MP2RAGE Image:An MPRAGE MRI sequence consists of a 180 degree inversion pulse followed by successive rapidly acquired gradient echoes obtained at short TEs and small flip angles.
Similarly, an MP2RAGE sequence uses two TURBO-FLASH GRE readouts between each inversion pulse. By combining image data from the 1st and 2nd readouts, T2* and B_1 inhomogeneity effects can be largely cancelled out, resulting in a strongly T1-weighted image with superior gray matter to white matter contrast. In this thesis, the MP2RAGE UNI image is used for full brain segmentation due to its superior T1 weighting and intrinsic uniformity. MP2RAGE images were acquired at 1mm resolution for better segmentation quality.



(a) T1 image

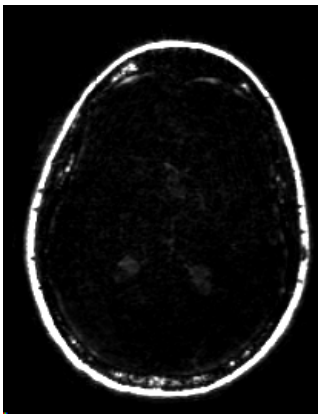


(b) T2 image

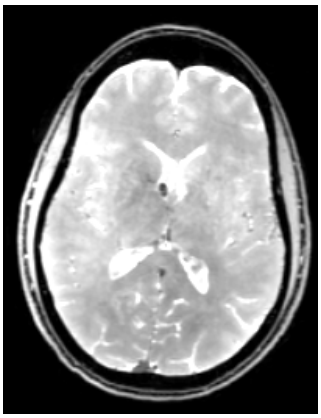


(c) Proton Density Image

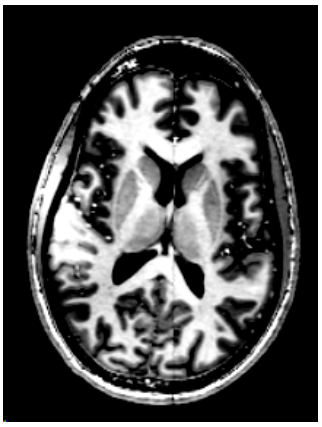
Figure 2.2: MRI Acquisition



(a) Fat Image



(b) Water Image



(c) MP2RAGE UNI image

Figure 2.3: MRI acquisition

Since, the 7T images are prone to artifacts and image intensity shading it is imperative that we have a consistent ground truth model which is later on used for training purposes in the deep learning approach mentioned in detail in chapter-3. Any errors in manual segmentation would reflect in the automated process and ultimately lead to incorrect SAR estimation. The processing of these data therefore includes several steps, which will be described in chapter-3.

2.3. DEEP LEARNING BASICS

The earlier chapters of this thesis describe how conventional image processing segmentation algorithms are time- consuming and cumbersome. With the onset of influx of data in everyday lives these manual and semi-automatic segmentations could take a long time for processing making them extremely time inefficient. Hence, it is now required to shift to artificial intelligence(either deep learning or machine learning) based approaches to handle a large amount of data. These algorithms are time efficient and can handle large amounts of data(images) efficiently.

Deep Learning is a subset of machine learning where it consists of algorithms that permit software to train itself to perform tasks like speech and image recognition.

Machine Learning is a subset of artificial intelligence that allows machines to perform tasks and improve on tasks based on prior statistical information.

Artificial Intelligence is a technique that allows machines to mimic human behavior or tasks using logic rules, if-else rules and decision trees.

The image below gives a clear depiction of the differences between the three

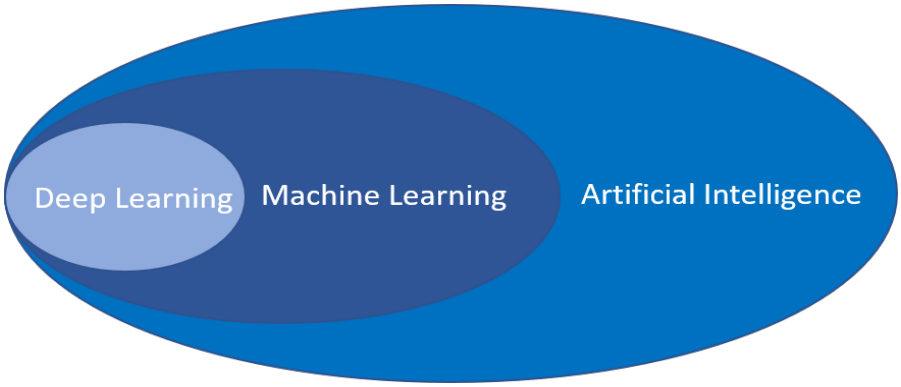


Figure 2.4: Deep Learning, Machine Learning and Artificial Intelligence

Deep learning networks use neural network architectures and hence are often referred to as Deep Neural networks(DNN). The network is called a "deep" neural network because it refers to the depth of the network, i.e., the number of hidden layers in the network. The amount of these "hidden" layers can vary from 2-3 in very simple networks to about 150 in very complex networks.
Deep learning models are trained using large datasets of labeled data and the network learns directly from the data without the need to extract features manually.

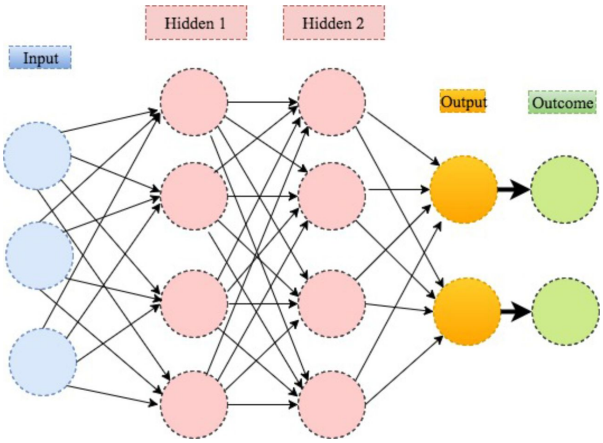


Figure 2.5: Deep Learning Network. [32]

The above figure depicts the complexity of the neural network which is organized in layers consisting of interconnected nodes. The network can have very few or a large number of hidden layers.
The deep learning network used in this thesis is the Convolutional Neural Network(CNN) which consists of a bunch of convolutional layers and de-convolutional layers. A CNN uses a convolution operator to convolve the features it has learned with the input. The input can be a 2D slice or a 3D volume(a stack of 2D slices).
CNN is used in this thesis because it eliminates the need for extracting features manually, so we do not have to identify features and then classify images. The network directly extracts features from the convolutional layers. The features are trained during the training process in which the network trains on a sample of images. This automated feature extraction makes deep learning models highly accurate for computer vision tasks.
CNNs detect the features automatically by using the many hidden layers that the network posseses. Every hidden layer increases the complexity of the learned image features. For example, the first hidden layer could learn how to detect edges, and the last learns how to detect more complex shapes specifically catered to the shape of the object we are trying to recognize.
Some common terms that are used in CNNs are explained below:

- **Convolution operation**: In mathematics, a convolution operation can be defined as a mathematical operation between two functions, that produces a third function and depicts how the shape of one function will change by the other function. As described earlier, the convolution operation convolves the input image with a set of convolution filters that extract features.

- **Receptive Field**: A receptive field can be defined as the area in the input volume that a filter is taking into consideration.
- **Max pooling operation**: A max pooling operator is like a decimation operator that down-samples the size of the feature map so that we have reduced resolution.
- **Transposed Convolution**: A transposed convolution is the inverse operation of a convolution operation, i.e., it can reconstruct the input signal before convolution. In deep learning terms, it is used to up-sample an image. However, at higher levels, the input volume is a low-resolution image and the output volume is a high-resolution image.
- **Activation Function**: An activation function in a CNN defines the output of a given node for a set of inputs of a particular node. The activation function is a node that is put at the end of or in between Neural Networks. In biological inspired neural networks they help to decide if the neuron would fire or not. We have different types of activation functions like sigmoid, Rectified Linear Unit(ReLU), Exponential Linear Unit(eLU), softmax, etc.
- **Overfitting**: Overfitting occurs in a CNN when the network performs well for training data but performs poorly on the test set. Essentially, the network has not learned anything and when a new test set is presented it cannot perform the task. In this case, the training accuracy will be much higher than the validation dataset. As can be seen in 2.6 It is discussed in detail in chapter-3, where we encounter the problem of overfitting and try to solve it. We can address the overfitting problem in the following ways:
  - Add more data
  - Use data augmentation
  - Use regularization like dropout layers and L1, L2 norm
  - The complexity of the architecture of the network can be decreased to solve the problem of overfitting.
- **Underfitting**:Under fitting occurs when the model performs poorly overall, i.e., the validation accuracy is much higher than the training accuracy. Though, in most networks overfitting remains a major problem, underfitting can arise if the model is less complex, so adding more layers or increasing the network complexity may reduce under fitting. Too many dropout layers can also cause under fitting.

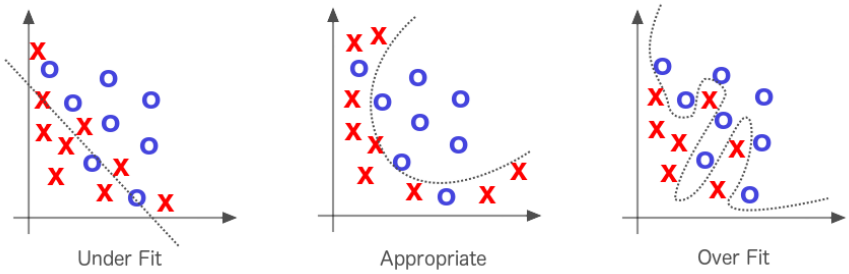


Figure 2.6: Underfitting, Correct Fit and overfitting[33]

2.3.1. DEEP LEARNING IN MEDICAL IMAGING

Manual image processing algorithms for segmentation are time-consuming and laborious, hence we need new techniques that can learn the features of the segmented images and predict segmentation on a completely new dataset. This is where deep learning algorithms would greatly benefit in medical imaging or image processing, in general.

One of the main applications in which deep learning is used in medical imaging is for identifying abnormalities in a perfectly healthy tissue or organ. Automatic image analysis algorithms based on machine and deep learning are the future of image processing. They would aid in quick diagnosis of diseases and would

facilitate efficient and robust tools. Applications of deep learning in healthcare covers a broad range of problems ranging from cancer screening and disease monitoring to personalized treatment suggestions. In this thesis, we mainly focus on the segmentation problem using deep learning approaches.

For medical image segmentation deep learning is extensively used these days, whether it is to detect tumor or in dosimetery studies. A very common deep learning based network that is used for medical image segmentation is the U-net[22]. The main idea of a U-net is to supplement a contracting network by successive layers. In this architecture, pooling operators are replaced by up-sampling operators. This, in turn increases the resolution of the output. To localize, high-resolution features from the contracting path are combined with the up-sampled output. A successive convolution layer can then learn to assemble a more precise output based on this information. One major feature of the U-net is that in the expansive path there is a large number of feature channels which allow the network to send information to higher resolution layers. This makes sure that the expansive and contraction path are almost symmetric and yields a U-shaped architecture and hence is called a U-net.

The basic architecture of U-net is illustrated in the below figure:

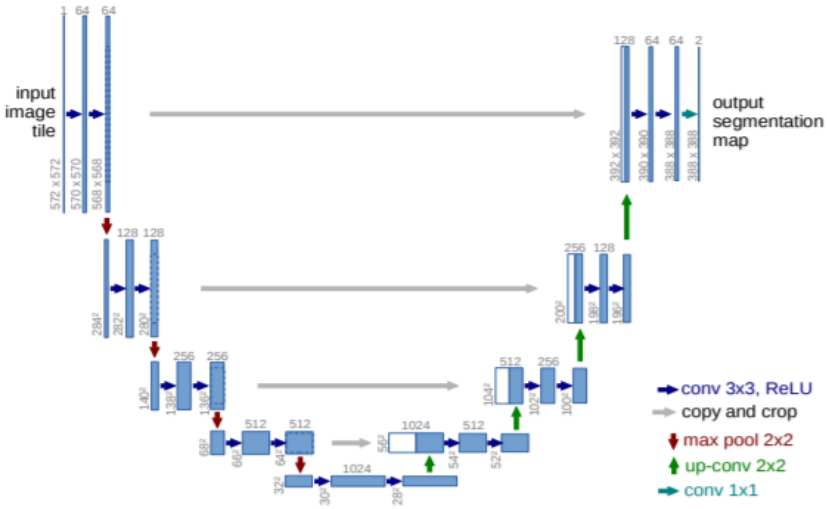


Figure 2.7: U-net Architecture[22]

The U-net architecture is like a typical CNN architecture. It has repeated convolution layers, each followed by an activation function. A max pooling layer is used for down-sampling. At the end of each down-sampling layer, we double the number of feature channels. Every step in the expansive path consists of an up-sampling of the feature map followed by a convolution ("up-convolution") that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and convolutions, each followed by an activation function. The cropping is necessary due to the loss of border pixels in every convolution. At the final layer convolution is used to map each feature vector to the desired number of classes. Eventhough U-net is primarily used for medical image segmentation it has some drawbacks and sometimes may not predict accurate results when the number of tissues to be segmented is large. An adaptation of the U-net is described in the next sub-section.

2.3.2. FORKNET ARCHITECTURE

Eventhough, for most image segmentation tasks U-net is the gold standard it has some de-merits. Some of these de-merits are stated below:

- To have good segmentation results, the size of the U-net must be comparable with the size of the features and its surroundings.
- U-net consists of numerous layers so it is not computationally efficient
- It can be very specific for a particular task

- It cannot handle multiple semantic segmentation and the loss function blows up when the number of classes(N) is large.

A novel CNN architecture presented in [21] solves this problem of semantic segmentation and can incorporate a large number of classes. This architecture is known as ForkNet and is based on U-net. In this architecture, for each tissue label there is a separate decoder path. So, for each tissue label there is one decoder path. In the author's research, there are 13 tissues,i.e, N=13. The MRI dataset that they have utilized is 3T dataset while we have to work with 7T dataset. It is of immense importance that the network takes into account the image inhomogenities present in the 7T images. For this purpose, the network must be retrained on 7T images. In the current thesis, the above ForkNet architecture was modified and retrained on 7T images from scratch. More convolution layers were added to the network to extract more features. Since, we have already established in Chapter-2 that we only need 6 tissues to correctly estimate SAR values. The tissues segmented in Chapter-3 are fed as labels into the new network along with the 7T T1 synergy image.

Figure 2.8 illustrates the ForkNet architecture for N=2 tissues:

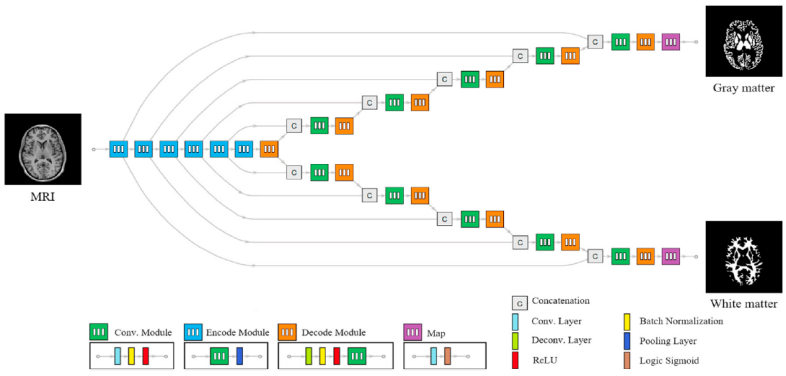


Figure 2.8: ForkNet Architecture[21]

In ForkNet, the input is a 256x256 grayscale image and the output is also a 256x256 labelled tissue image. The tissue label image is binary. The loss function used in ForkNet is binary cross-entropy, an ADAM optimizer is used for minimizing the loss and maximizing the performance metric. As a performance metric, the authors have used dice coefficient and haussdorf distance. The batch size is 4, and the network is allowed to run for 50 epochs. However, there are certain limitations associated with the deep learning approach. The quality of ground truth segmentations would never be completely accurate. There are always errors in the ground truth, these errors then also travel to the output of the network. Since, the dataset of MRI images is 3T, many non brain regions are blurred and have low SNR and could not be segmented properly. Since MRI is prone to a large set of inhomogeneities due to scanning parameters, it makes the algorithm not robust for field and bias inhomogeneities.

3

METHODOLOGY

This chapter describes in detail the methodology adopted for the current thesis. Section 3.1 describes the tissue reduction study. Section 3.2 describes the image processing and segmentation pipeline. Lastly, section 3.3 describes the deep learning aspect of the thesis.

3.1. CHRONOLOGY

This thesis aims at accurately estimating the RF power SAR deposition in 7T MR scanners. For this, we use high contrast images from the MR Scanners to create anatomical body models from the volunteers. High contrast images are used to facilitate easier manual segmentation and establish a reliable ground truth efficiently. A list of dielectric and physical properties of tissues is created. These properties are then mapped onto the corresponding physical properties to perform numerical simulations. SAR maps and B_1^+ maps are computed using (XFdtd 7.4, Remcom inc., State College, USA). A Deep Learning method is then developed to infer a similar anatomical model from T1 images. The models are then compared to the ground truth model. SAR power deposition is the main parameter to be compared.

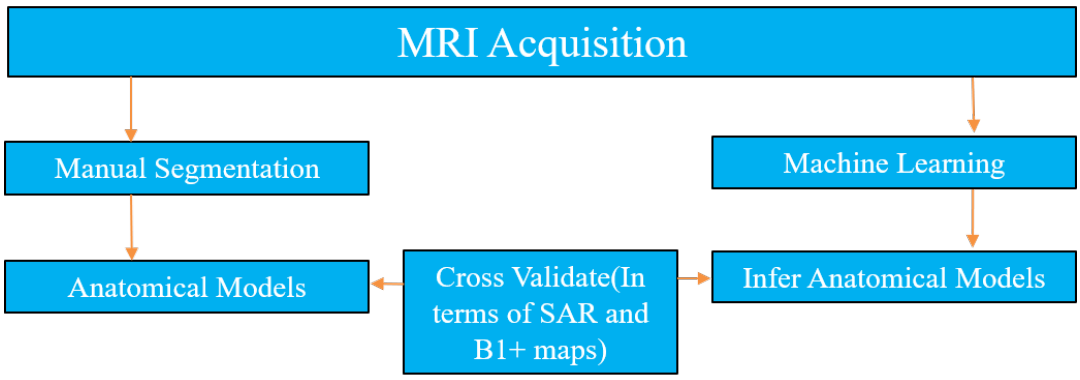


Figure 3.1: Methodology

The methodology from the above image is explained here:

1. 7T images are acquired from the MRI scanner. The multi-contrast images that are acquired are: T1 weighted, T2 weighted, MP2RAGE, Dixon and Proton Density weighted image.
2. The images are manually and semi-automatically segmented and high contrast anatomical models are created from it.
3. A Deep Learning algorithm is used to segment tissues automatically
4. Cross-validation between the deep learning method and the ground truth is done using SAR maps as a parameter.

3.2. TISSUE REDUCTION STUDY

As already described in literature models in chapter-2, tissue reduction has to be performed to make the segmentation process easier. One of the goals of this thesis is to determine the number of tissues required for a robust SAR estimation. Many studies have been conducted to predict the appropriate number of tissues that need to be segmented in order to reach a realistic SAR estimation. Some of these studies have used a three tissue model, while some have used a thirteen tissue model. However, most of these studies have been conducted on 3T images. Hence, it is unclear whether the results derived from these studies would also predict 7T SAR predictions based on 7T image data. In this thesis, Duke model from the Virtual Family was used to study the effect of a reduced body model on SAR values as compared to the full body model. The results were compared based on 10g averaged SAR maps and dielectric properties. A quadrature birdcage coil was used for this purpose. The results were computed using (XFdtd 7.4, Remcom inc., State College, USA).

The Duke model consists of 47 tissues and 47 tissues for the head and torso.

A scatter plot of 47 different tissues of the head and torso is represented in the below figure. It is clear from the figure, that the majority of the tissues can be classified as muscle, but CSF, which has the highest conductivity needs to be segmented. Fat is essential for segmentation as it behaves like an isolator, defining induced current pathways, SAR is an integral effect, changes in one location may influence SAR in another area.

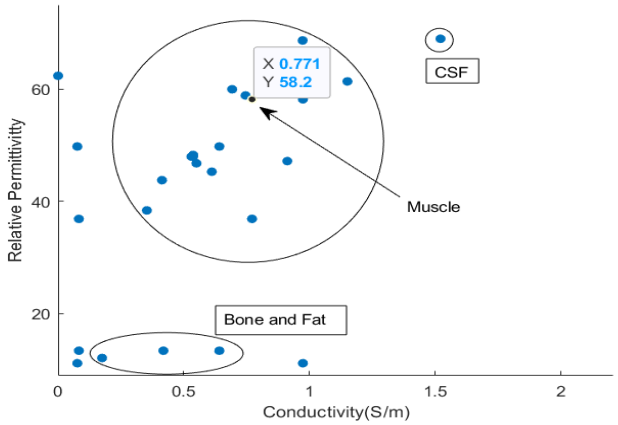


Figure 3.2: Dieletric Properties of Tissues

Some body fluids like CSF have high conductivity and permitivitty and contribute towards estimation of SAR values. Since the conductivity of CSF is large, RF power deposition mainly occurs in such tissues. The fat distribution in the body has a profound effect on wave propagation and eddy current pathways in the body. Most other tissues can be classified as muscle(water) in the human body.

One of the sub-goals of this thesis is to determine the appropriate number of tissues to be taken into consideration while creating the numerical body models. Earlier studies have proved that it is possible to reduce the number of tissues to get accurate SAR estimates. We would take a sequential approach to find out the number of tissues needed for segmentation. We would compare the reduced duke models with the full duke models to draw a conclusion. We would consider the following duke reductions: two tissues reduction, three tissues reduction, five tissues reduction, six tissues reduction and thirteen tissues reduction.

To test the two tissue model, the duke model was reduced from 47 tissues to just 2 tissues. This is the easiest model when the ease of segmenting tissues is considered. The clustering of tissues was performed based on similar dielectric properties of the tissues. Tissues having high conductivity like CSF, need to be segmented to have correct SAR estimates. If these tissues are grouped with tissues having lower conductivity, SAR maybe affected and unrealistic hotspots maybe obtained. The two tissue model did not yield realistic SAR estimations in the case of 7T images.

Since, the two tissue model did not yield sufficient results, a three tissue model was segmented and simulated. The detailed Duke model was segmented into muscle, fat and brain. The brain was segmented as one

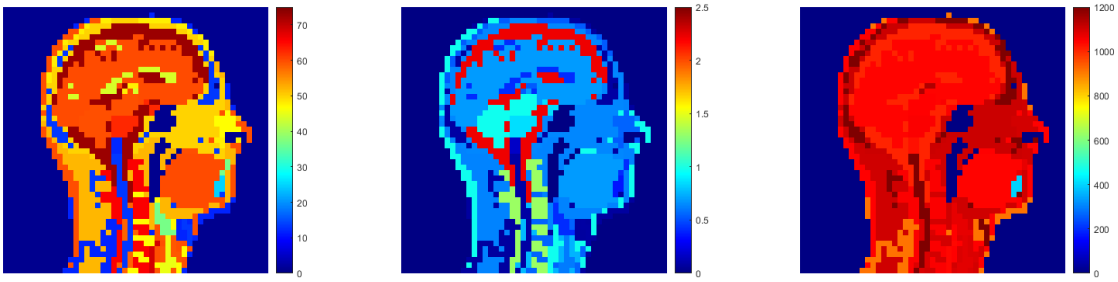
tissue and was not segmented further into grey matter, white matter and CSF. In this model, even though the brain was differentiated from the muscle. The conductivity of CSF, which is a brain tissue, is the highest, so segmenting CSF may lead to a realistic SAR estimate.

Since, the three tissue model improved the SAR estimations but still did not provide accurate estimates, a five tissue model was used to simulate the effect of these tissues on SAR hotspots. Now, the brain was segmented into white matter, grey matter and CSF. Grey matter and white matter typically have intermediate conducitivities and permitivitties, whereas CSF has a very high conductivity of 2.2 S/m. This makes CSF an extremely relevant tissue for SAR estimates. So, this tissue model consisted of grey matter, white matter, CSF ,fat and muscle. The SAR hotspots were quite similar to the full detailed Duke model. However, since there is no perfusion in the eye so UHF systems may lead to local heating near the eyes. Hence, incorporation of eyes in the model may lead to better SAR estimates.

The five tissue model performs quite well as compared to the two tissues and three tissues models. The SAR estimates derived from the five tissue model are quite comparable with the full detailed model, however since, the eye(vitreous humor) is a high conductive region, it also needs to be segmented. The six tissue model consists of grey matter, white matter, CSF, fat, muscle and eyes. The SAR estimations were very similar to the detailed Duke model.

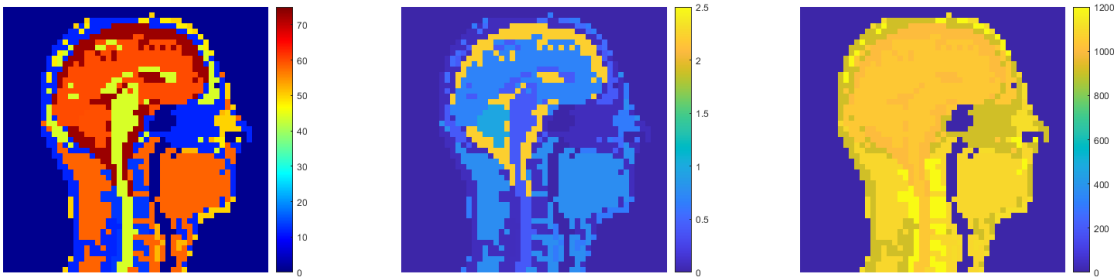
The six tissue model had similar hotspots as the full tissue model which consisted of 47 tissues. We wanted to further investigate the affect on SAR estimations when we increase the number of tissues. So, in this model, the tissues were increased to thirteen. This was motivated by [21] in which the authors segmented thirteen tissues using a deep learning based approach. The thirteen tissue model consisted of grey matter, white matter, CSF , fat, muscle, dura, eyes, skin, skull, cerebellum, bone(cortical), bone(cancellous) and blood vessels. This model yielded similar estimates as the six tissue model and the manual segmentation of thirteen tissues is a very cumbersome task in 7T as will be described later. Hence, it is not required to segment the head model into thirteen tissues. Since, most of the extra tissues segmented have a minor influence on SAR distribution as they have very similar properties with the tissues already segmented. For instance, the properties of cerebellum and grey matter are similar and they can be combined together and labelled as grey matter.
So, the 47 tissues DUKE model can essentially be reduced to just six tissues without affecting major SAR hotspots.
The tissues can be combined in many ways, in many studies k-means clustering is used to group different tissues together.[34] However, it may lead to mis-classficiations in tissues which may seriously reduce SAR estimates. In this thesis, we grouped the tissues based on their dielectric properties and ease of segmentation. Tissues having similar dielectric properties were grouped together and classified as a single tissue.

For the Duke Model:



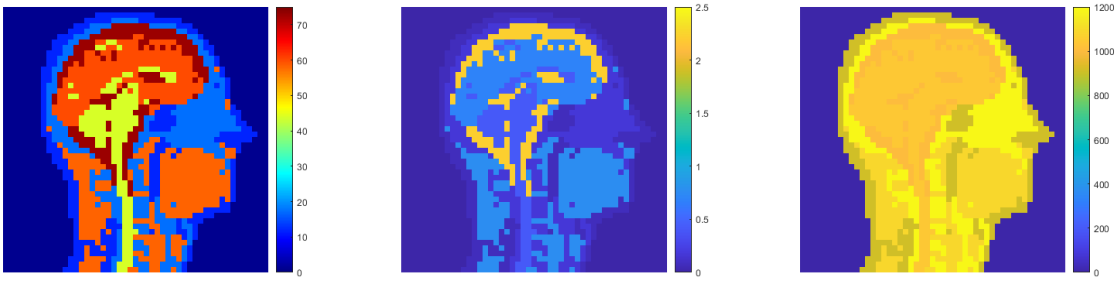
(a) Relative Permittivity (b) Conductivity (c) Density

Figure 3.3: Dielectric and Physical Properties for Full Duke



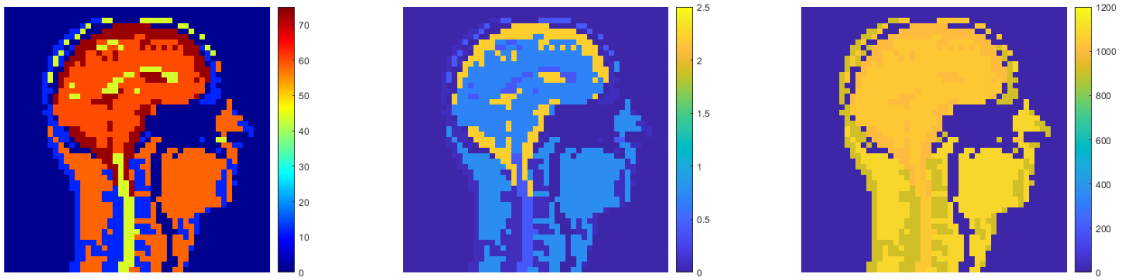
(a) Relative Permittivity (b) Conductivity (c) Density

Figure 3.4: Dielectric and Physical Properties for thirteen tissue Duke



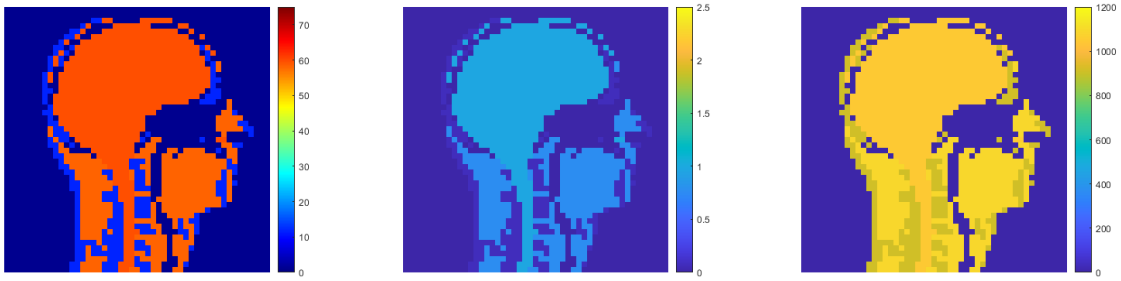
(a) Relative Permittivity (b) Conductivity (c) Density

Figure 3.5: Dielectric and Physical Properties for six tissue Duke



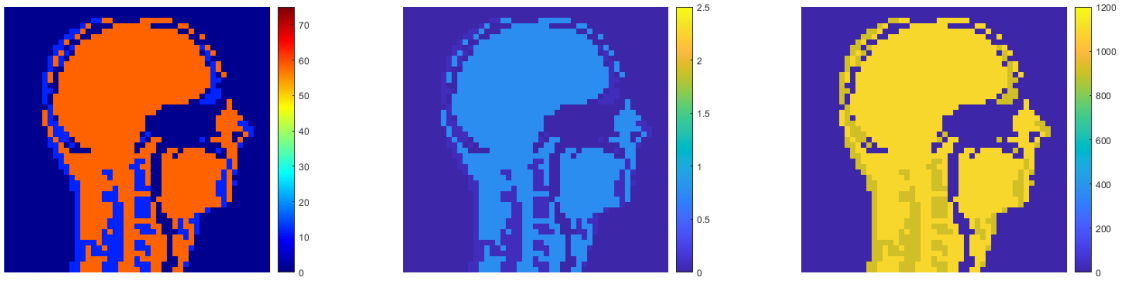
(a) Relative Permittivity (b) Conductivity (c) Density

Figure 3.6: Dielectric and Physical Properties for five tissue Duke



(a) Relative Permittivity (b) Conductivity (c) Density

Figure 3.7: Dielectric and Physical Properties for three tissue Duke



(a) Relative Permittivity (b) Conductivity (c) Density

Figure 3.8: Dielectric and Physical Properties for two tissue Duke

Figures 3.3 to 3.8 depict the dielectric and physical properties of the duke model and denote the different segments as tissues.

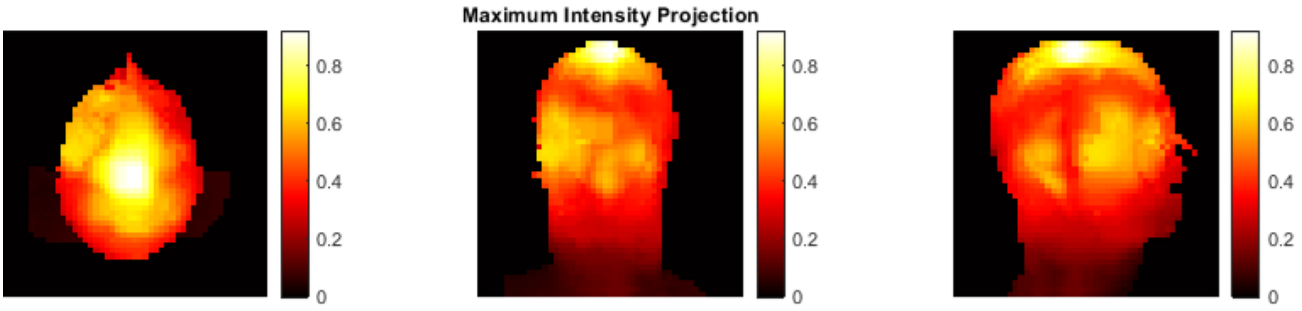


Figure 3.9: 10 g average SAR maps for full model

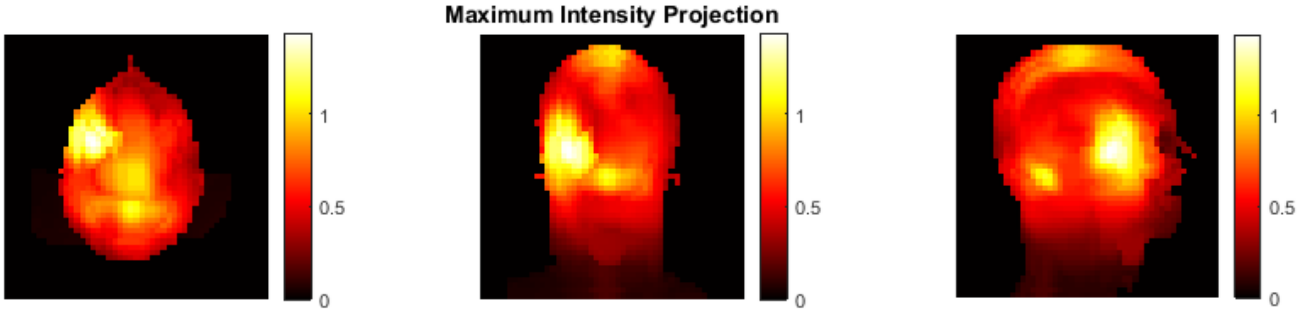


Figure 3.10: 10g average SAR maps for 13 tissue model

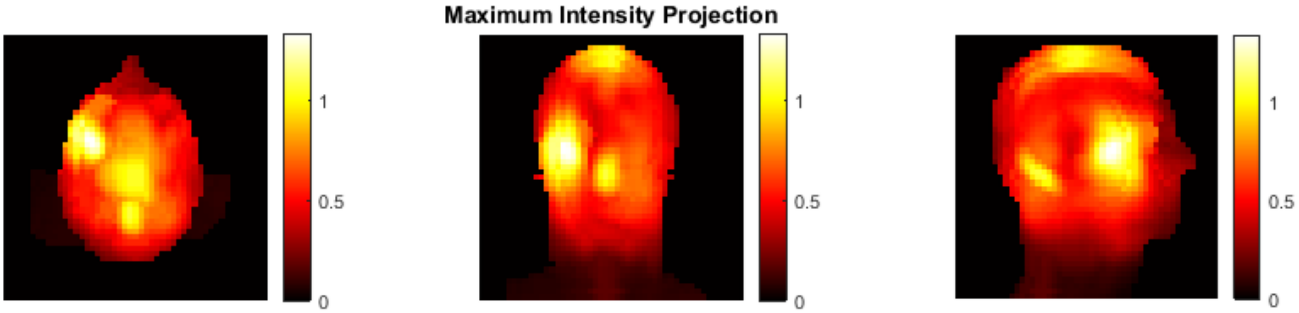


Figure 3.11: 10 g average SAR maps for 6 tissue model

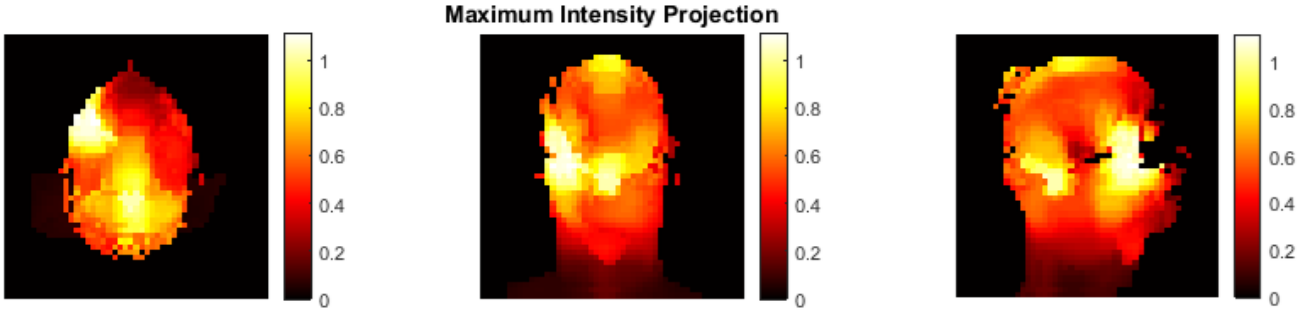


Figure 3.12: 10g average SAR maps for 5 tissue model

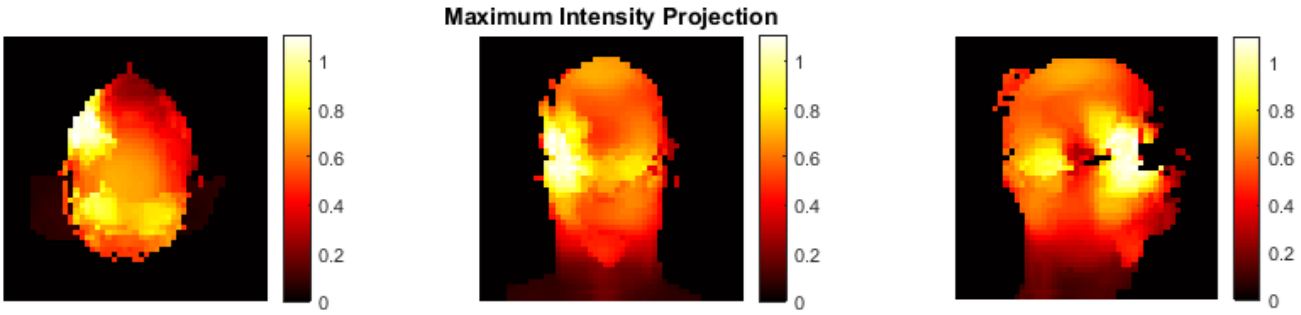


Figure 3.13: SAR Maps for 3 tissue model

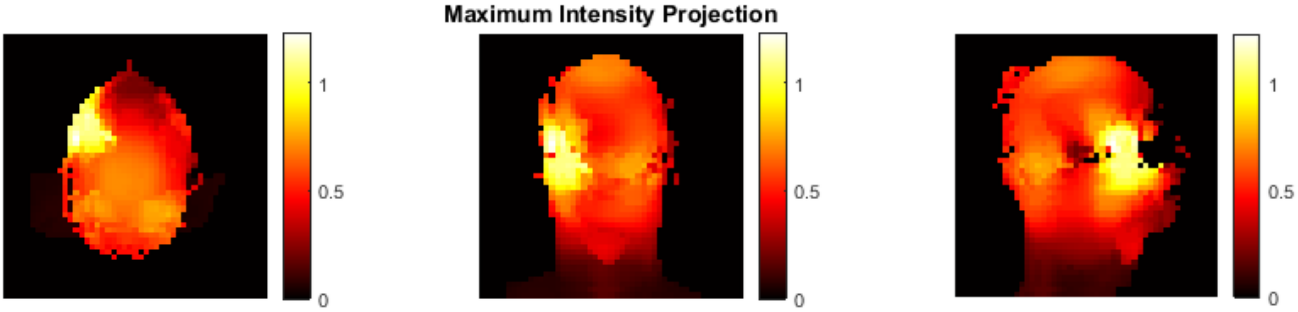


Figure 3.14: SAR Maps for 2 tissue model

After analysing the SAR maps of all the cases presented above it has been found out that the appropriate number of the number of tissues that have to be segmented is six. It is quite evident hat two and three tissue models deviate from the full model SAR values. Unlikely hotspots appear in the two tissue and three tissue model, this happens since the tissues have been clustered improperly. Hence, it was concluded that the two and three tissue models do not yield sufficient results and hence are not the optimal number of tissues that have to be segmented. There are a lot of similarities between the 5 tissue and 6 tissue model, the only difference between these models is that in the six tissue model, eyes ere also segmented along with the brain. Even though, the conductivity of eye(vitreous humor) is quite low, it is a good practice to include it into the tissue model as the eye tissue can't dissipate heat due to perforation, so sometimes hotspots can be formed near the eyes as in the case of full model SAR values depicted above. Segmenting the eyes is an easy process and is more time efficient, so including the eye tissue in the model does not hinder the ease of the image processing.

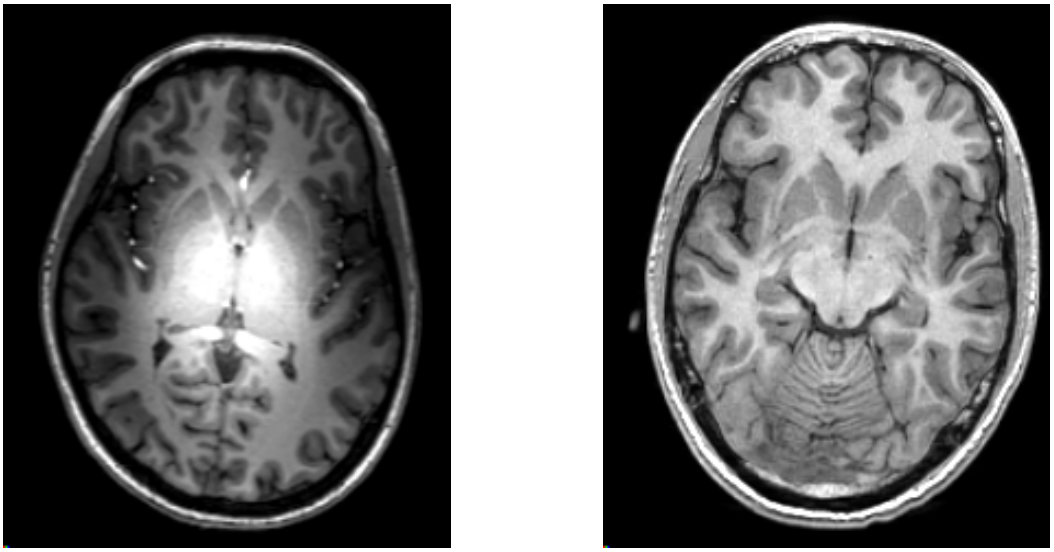
The full Duke body model was reduced from 47 tissues to just 6 tissues. In a study by Rashed, et al, considered a 13 tissue model, but in this case the six tissue and 13 tissue model generate similar results. Hence, segmenting the tissues into six distinct types is the optimal approach.

3.3. IMAGE PROCESSING AND SEGMENTATION

Various sequences that have been acquired for the current thesis have been explained in chapter-2. Since, these images are acquired using a 7T scanner they are subjected to image inhomogenities that need to be corrected before we start segmenting these images. Accurate ground truth models have to be created from the multi-contrast anatomical images. These images are then segmented using manual and semi automatic processes which are explained in much detail in the following section. Since, these images have to be segmented, pre-processing of these images is required before we can segment them. Image inhomogenity correction and registration is performed before the images are segmented into various tissues.

3.3.1. IMAGE INHOMOGENEITY CORRECTION

The 7T images are majorly corrupted due to a) Field inhomogenity and b) ghosting artifacts. In ultra high field MRI scanners, the magnetic field strength is very high. Due to this increase in the B_1 field, a lot of inhomogeneties are prominent in 7T images. Image inhomogeneity is also present in 3T images but it's higher in 7T images which remains a major bottleneck for any image processing algorithm. As most segmentation methods are based on some form of thresholding function, the segments would be mis-classified if the image inhomogenities are not addressed properly. These segmentation errors may also affect SAR simulations, which would lead to an incorrect estimation of SAR distribution.



(a) Intensity shading in a 7T image

(b) Intensity shading in a 3T image

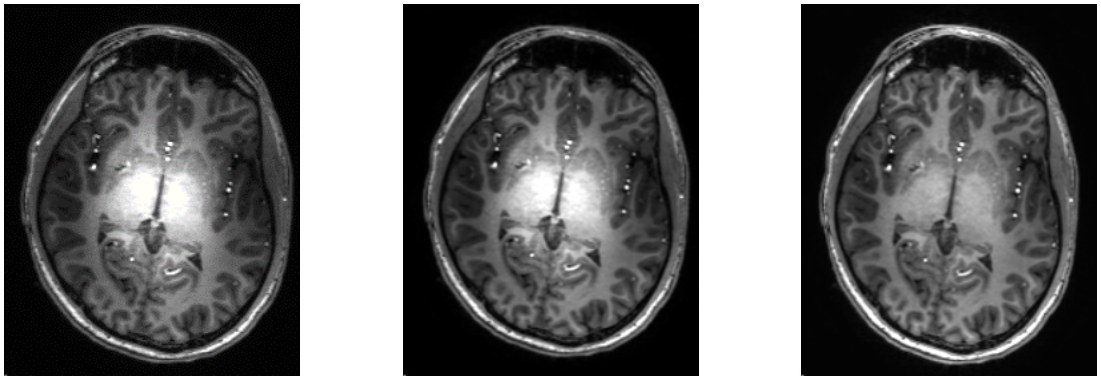
Figure 3.15: Image inhomogenities

A comparison of the image inhomogenities in 7T and 3T images is presented in . It is clear from the above figure that image inhomogeneity is a big problem in 7T images.

Many algorithms like N3 and N4 corrections have been used in the past to correct image inhomogenities.[35]. However, these methods fail on the 7T images as there is still a lot of residual inhomogenities left. Background removal must be done in a proper way in order for these algorithms to work. Since there is background noise present in the images it is difficult for these methods to perform in an optimum manner for 7T images. These methods are computationally expensive. Inhomogenties are still left in the image and N4 is not able to correct that. So, another method has to be deployed. The comparison is illustrated in fig 3.15.

A novel method to solve the image homogeneity problem is implemented in [36]. In this method a B_1^+ map is used to account for the intensity shading. The DREAM method was used for multi-slice B_1^+ mapping. Receive sensitivity B_1^+ maps were obtained via the DREAM data as derived in:

$$\begin{aligned} I_{FID} &= M_0 B_1^- sin\beta sin^2\alpha \\ I_{STE} &= \frac{1}{2} M_0 B_1^- sin\beta sin^2\alpha \\ M_0 B_1^- &= \frac{(I_{FID} + 2I_{STE})}{sin\beta} \end{aligned} \tag{3.1}$$



(a) T1 image with image inhomogeneities (b) T1 image after N4 correction (c) Image inhomogeneties correction in 7T images

Figure 3.16: Image inhomogeniety in 7T images

3.3.2. REGISTRATION

Image registration is a process in which different images having different characterstics can be mapped onto one coordinate system. The difference in characteristics of the two images could be due to rotation, displacement or shear. This is mainly used in computer vision, for instance, to register images from different modalities(Eg, a CT image registered with an (MRI image) or inter patient registration,i.e., registering images from different patients onto one main image. Sometimes, it is also used to register images of one patient acquired in different points of time. In this thesis we need to correct for subject motion in between acquisitions of different MR contrasts, when the displacement was more than 0.5 mm.

Rigid Registration:In Rigid Registration, two images having very different coordinate systems and orientations are mapped onto one coordinate system. This happens by calculating a translation and rotation which places the images or certain parts of the images at the desired place. This process is known as rigid registration. Most of the approaches to find rigid registration work on trying to minimize the distance between two pixels of corresponding images.

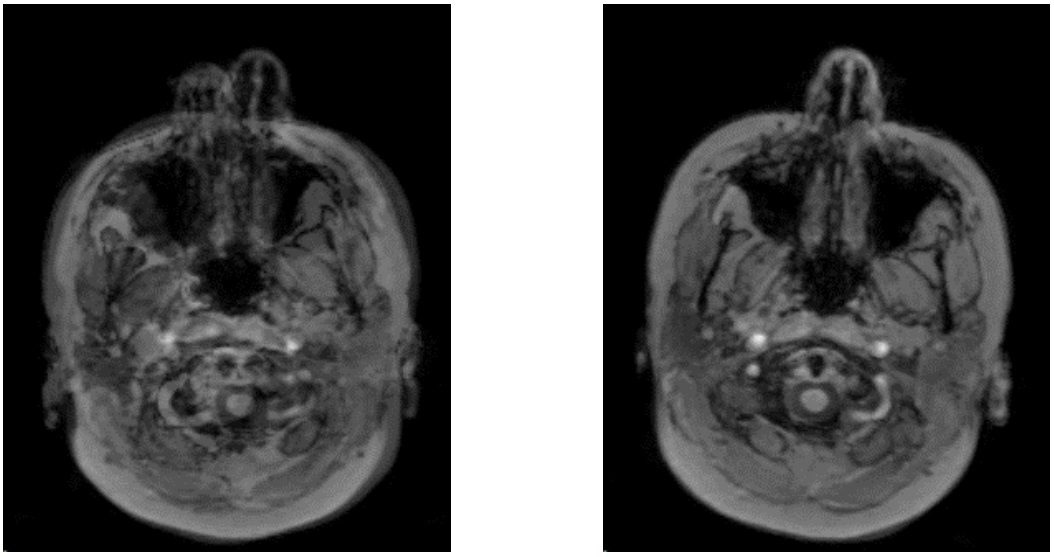
Non Rigid Registration:In non rigid registration, B-splines are used to parametrize a free form deformation field(FFD). This is a complex registration problem as opposed to a rigid registration because of a higher-dimensional parameter space.

The following algorithm was used for image registration:

- MP2RAGE INV2 image was registered on dixon in-phase image,i.e., the in-phase image was fixed while the MP2RAGE INV2 image was moving. This was done using the elastix module in 3D slicer.[14] This registration is feasible because MP2RAGE INV2 and the dixon in phase have very similar contrasts.

- The transform from the registration output of MP2RAGE INV2 image is used to register MP2RAGE UNI image.

Rigid registration is used in this case due to the following reasons: a)The patient is same, b)The modality is same and c)There are little changes in the brain shape or position within the skull due to the above two reasons. So, the rigid registration is more robust.



(a) In Phase image and MP2RAGE UNI image before registration

(b) In Phase image and MP2RAGE UNI image after registration

Figure 3.17: Registration

3.3.3. SEGMENTATION

As discussed in section 3.1, the target number of tissues for the segmentation process is 6. These tissues need to be segmented to make a patient specific-body model to correctly estimate the SAR value. The tissues taken into consideration are: White Matter(WM), Grey Matter(GM), Cerebrospinal Fluid(CSF), Fat, Muscle(Water) and Eyes. A multitude of softwares have been used for this manual and semi-automatic processes. These include: ITK-SNAP[12], FSL, 3-D Slicer[14], MevisLab[37] and MATLAB. Since we have multi-contrast data, we can use the data for different segmentations. In most studies, T1 image has been used for segmentation of the brain. T2 image can be used to segment CSF and eyes as they have improved contrast in the T2 weighted image. The dixon sequence can be used to segment fat and muscle.

Segmentation of the Brain: The brain is segmented into grey matter, white matter and CSF. Initially, the brain scans were acquired at 2mm resolution. However, it was difficult to segment these 2mm scans as the boundaries between segments were blurred. As mentioned earlier, in chapter-2 we consider the 1mm isotropic resolution for all images as there is less blurring along the edges of a 1mm image which makes the segmentation process more accurate.

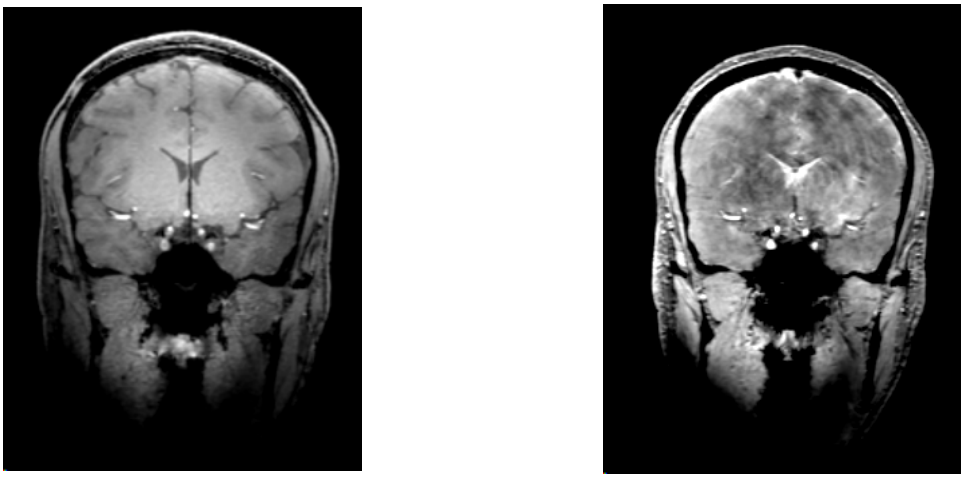
We use MP2RAGE images for segmenting the brain because they are automatically corrected for intensity during acquisition. The mask from the in-phase d MP2RAGE INV2 image is first corrected for image inhomogeneities using B_1 correction. N4ITK[35] is again used to correct for some residual inhomogeneities. MP2RAGE INV2 image is used because it has less background noise as compared to the UNI image and its contrast is very similar to a proton density image. [16]

The mask from the MP2RAGE INV2 image is then used on the MP2RAGE UNI image to obtain a brain mask that can be used in FSL's FAST segmentation toolbox to segment the brain into white matter, grey matter and CSF. Care was taken to manually correct the brain mask before feeding the data into FSL, to prevent segmentation errors around the borders of the brain. In some cases, manual correction was also needed after the FSL segmentation was obtained. Fixing the brain mask manually is a labour-some task and demands hours of your time.

Instead of using the MP2RAGE INV2 image for skull stripping, we now used the in phase image from the dixon. The in phase image has a thinner layer of skull which makes the skull stripping process quite convenient as illustrated in fig 3.12. The brain mask is then derived from the dixon in-phase image. It still needs to be manually corrected but the manual correction reduced significantly to just 10 minutes! The manual corrections are done using 3D-Slicer[14]. The mask is then applied to the MP2RAGE UNI image and then segmented using FAST. There are still some segmentation errors in CSF of about a 1mm voxel layer, but they

are acceptable as far as SAR simulations are concerned.

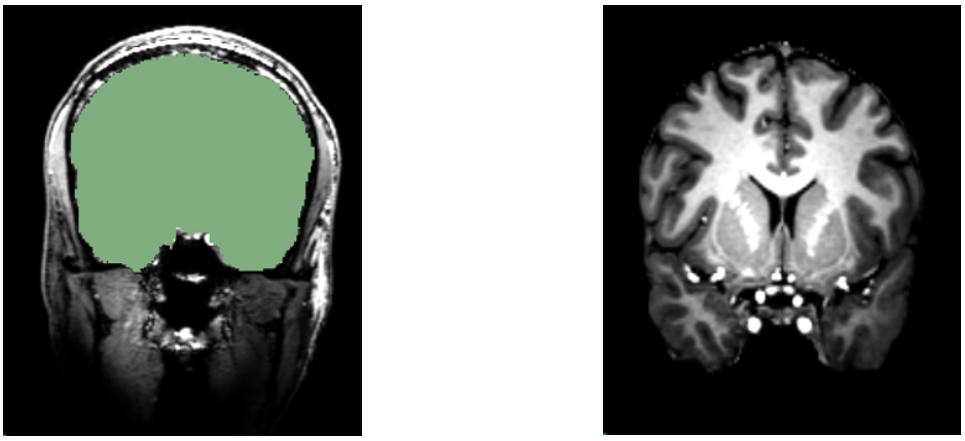
Brainstem is segmented using the MP2RAGE image. Fast Marching algorithm is used to segment the brainstem, it works in a similar fashion like a region growing method. A seed is selected in the region of interest and then it expands, segmenting the required area. The brain stem is later added with the white matter and CSF masks respectively.



(a) MP2RAGE INV2 image

(b) Dixon in-phase image

Figure 3.18: Comparison between MP2RAGE INV2 and Dixon in-phase image for brain masking



(a) Binary labelmap of the brain mask

(b) Brain mask , masked with the MP2RAGE UNI image

Figure 3.19: Brain Mask

The Brain segmentation pipeline is as follows:

- Dixon in phase image is first corrected for image inhomogeneities using B_1 correction as explained in the earlier section. N4ITK[35] is again used to correct for some residual inhomogeneities.
- Dixon in phase image data is used to separate the skull from the brain using BET Toolbox[38] in FSL(FMRIB, Oxford, UK).
- The mask so obtained was used to derive the brain from the MP2RAGE UNI image which has good contrast between the grey and white matter.
- The brain mask was then manually corrected using 3D slicer[14].
- The segmentation was performed using FSL FAST Segmentation toolbox.

Segmentation of Fat and Water: The dixon sequence is used to segment water and fat fractions. A simple thresholding technique can be used to segment water and fat, however, it is not as simple as it sounds. The dixon sequence is prone to a lot of ghosting artifacts which makes the segmentation process cumbersome. A mask needs to be used to get water and fat fractions.
For the masking process a proton density weighted image was used. For calculating the threshold, a poisson distribution-based minimum error thresholding function is used to minimize background noise. The acquisition protocol was improved in subsequent scans which also improved the TFE proton density weighted image. This image was then used as a mask to segment water and fat.
The in-phase dixon image is used as a mask which provides a larger signal in the neck but increases the background noise and also introduces the motion artifacts around the eyes. To solve this problem, background noise removal is done using connected components in MATLAB. Later on, this mask is manually edited in 3D-Slicer[14] before using on the water and fat images. Fat and water were segmented by classifying fat and water voxels. A voxel is labeled as fat if its intensity was greater than the fat voxel in the fat segment and a voxel was labeled as fat if its intensity was greater than the water voxel in the water segment. That translates into the following:

$$WF = \frac{Water}{Water + Fat} \tag{3.2}$$

Where, WF is Water Fraction

$$FF = \frac{Fat}{Water + Fat} \tag{3.3}$$

Where, FF is fat fraction
Segmentation of Eyes: Eyes were segmented using ITK SNAP. A threshold was carefully selected that separated the eyes from the background. Then a region growing algorithm using deformable snakes was used to segment the eyes.[39].
Now, that all the six tissues have been segmented using the process explained above, we need to create body models from these segmentations to estimate SAR. We need to combine the segments while taking account of the priority of different segments to preserve the relevant tissues.
These segments are combined in a single integer label map which consists of all these individual segments. It is saved as a nifti file and exported to (XFdtd 7.4, Remcom inc., State College, USA) for calculating SAR distributions. SAR estimates for high contrast multi resolution images are obtained from these segmentations. The ground truth segmentations created from the multi-contrast dataset are used to train the deep learning network which is discussed in detail in the next section.

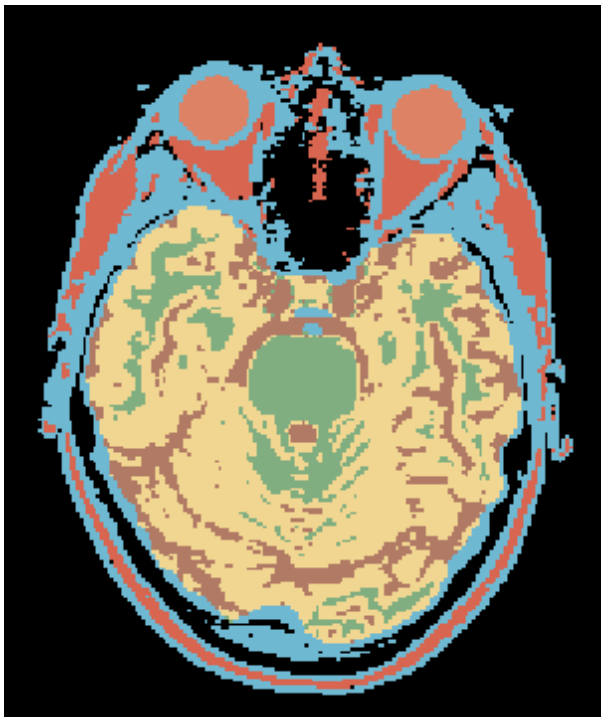


Figure 3.20: Body Model for a volunteer

3.4. DEEP LEARNING ARCHITECTURE

In the previous sections, multi-contrast images were used to create the ground truth for deep learning training purposes. The ground truth labels obtained from the high contrast images serve as an input to the network. A modified version of the U-net, called the ForkNet which is described in detail in chapter-2 is used for segmentation, however, it is trained on 3T images and needs to be modified and trained again for the current problem of the thesis.
There are numerous problems associated with the ForkNet, as illustrated in the previous section. To eliminate those problems, data was acquired at 7T scanner. The 7T images have high SNR and thus can image non brain regions better in comparison to 3T images. Since, the 7T images are prone to magnetic field bias, the network would be trained to incorporate the field bias. The architecture is adapted from ForkNet's[21] architecture and trained on 7T T1 weighted-images. Certain layers and parameters from the original architecture were changed to tune the network to give optimal results.

Initially, we used the binary cross-entropy loss function. The cross-entropy loss can be defined as:

$$CE = -\sum_{i=1}^{C} t_i log(S_i) \tag{3.4}$$

Where, t_i is the ground truth data, S_i is the output of the CNN for each class. An activation function(either softmax or sigmoid) is applied to the scores before computing the loss. Sigmoid is used incase the output is binary, otherwise softmax is used.
In a binary problem, where C=2, i.e., foreground and background, the cross-entropy function is called a binary cross-entropy function and can be defined as:

$$CE = -\sum_{i=1}^{C=2} t_i log(S_i) = -t_1 log(S_1) - (1-t_1) log(1-S_1) \tag{3.5}$$

However, the binary cross-entropy loss function cannot be used in the problem defined in the current thesis because the data set is imbalanced.

The concept of an imbalanced dataset is accurately explained using the following example. Suppose, we have two classes, C0 and C1. C1 has a gaussian distribution of mean 1 and variance 0, while C0 has a gaussian distribution of mean 2 and variance 2. Moreover, let us assume that class C0 represents 90% of the dataset and class C1 represents the remaining 10%. A better illustration is depicted in Fig.1.5

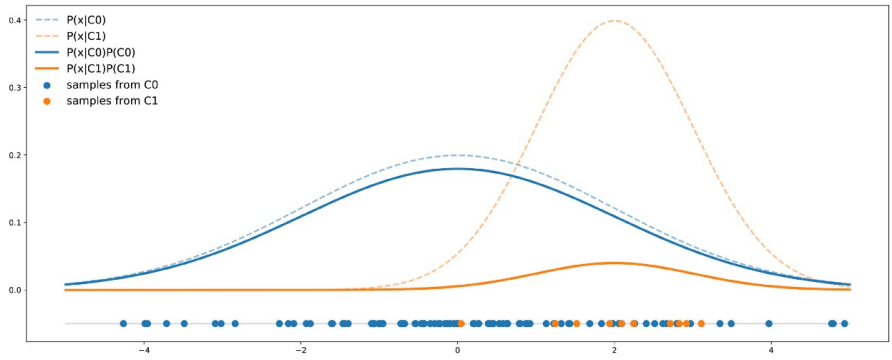


Figure 3.21: Illustration for data imbalance

Since, the curve of C0 is always above C1, at any given point, the probability that this point is from C0 is always more. In terms of images it essentially translates into, if the label is very small as compared to the image, i.e., the foreground constitutes only 1% of the whole image, then, the probability that the network will predict only 0s(background) is high. Since, it is only predicting the black pixels, the predicted image is a plain black image, even though the accuracy is very high.
The Dice loss is more robust to class imbalance and is presented in the current thesis. The Dice coefficient is defined as:

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \tag{3.6}$$

The Dice coefficient is nothing but, 2 × the Area of overlap divided by the total number of pixels in both images. Suppose the combined number of pixels in both images is 200 and area of overlap is 0, so foreground=(2× Area of Overlap)/(total pixels combined) = 0/200 = 0, while for the background are of overlap is 95, so (2 × Area of Overlap)/(total pixels combined) = 95×2/200 = 0.95, in the case of accuracy this is interpreted as accuracy while it is only predicting blank images. So, Dice = (foreground + Background)/2 = (0%+95%)/2 = 47.5% The accuracy curve for the binary cross-entropy loss is illustrated in Fig 3.22

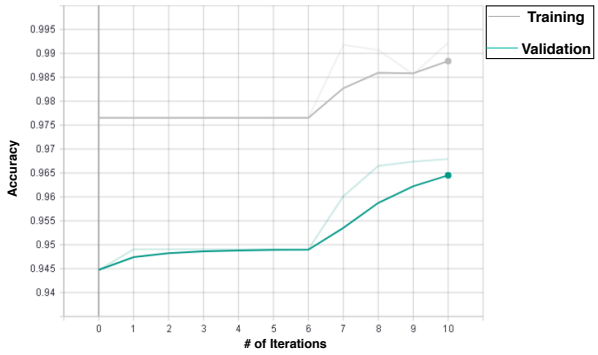


Figure 3.22: Accuracy Curve

As we can observe from the above figure that accuracy is high even at the start of the training, means that it is only going to predict a plain black image which we can see in the next figure. The green curve is the validation curve while the grey curve is the training curve. To address the problem of class imbalance we use the Dice coefficient as a loss function. However, one problem with the Dice coefficient is that it is not differentiable, so we have to convert it into differentiable form to use it as a loss function. So, we transform the Dice coefficient by subtracting 1 from it. The Dice loss now becomes:

$$Dice = 1 - \frac{2|A \cap B|}{|A| + |B|} \tag{3.7}$$

Dice coefficient is used as a performance metric for the network. The goal is to have minimum difference between training and validation datasets.

3.4.1. TRAINING

We use the data acquired from 8 volunteers for training purposes. The six different tissues were already segmented using the segmentation pipeline explained in chapter-2. The input is a 192*256 axial image along with 6 different binary labels of different tissues. Initial tests were performed for N=2, i.e., grey matter and white matter. A Basic architecture for six tissue segmentation is illustrated in Fig. 3.23.

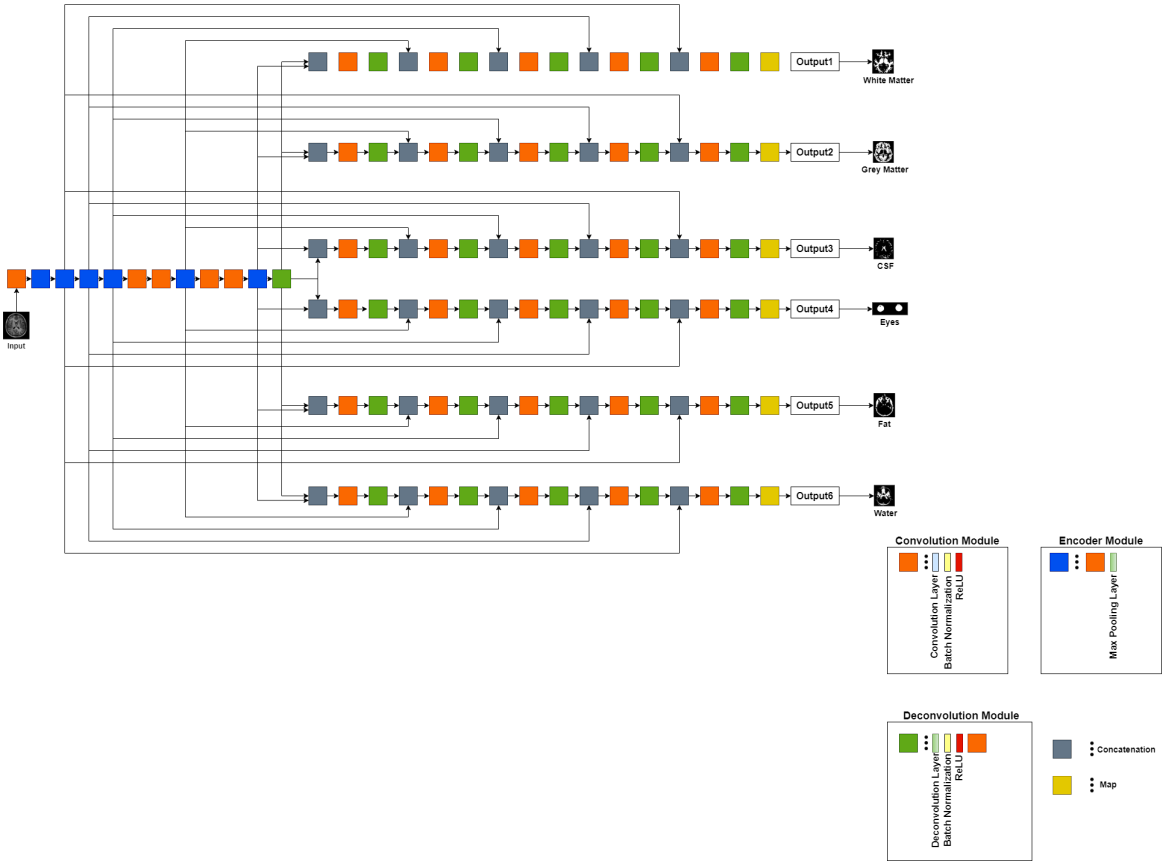
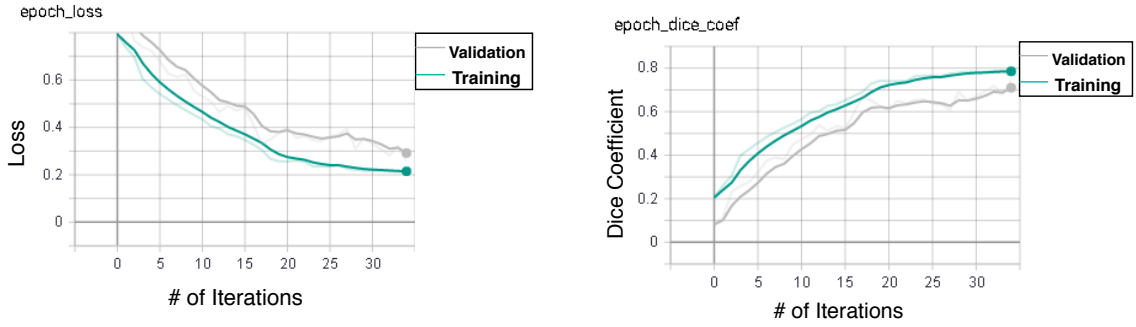


Figure 3.23: Adapted ForkNet architecture

This is the final architecture that was used to segment six tissues. The Adapted ForkNet was used in this thesis. Additional convolution layers were added to the pre-existing ForkNet to extract more features from the images. Since this network was slightly different than the ForkNet, we call it Adapted ForkNet. Initially, the ForkNet architecture was deployed for segmentation. The ForkNet architecture did not yield sufficient results for the 7T data even after being re-trained. Overfitting was described in chapter-2. The problem of overfitting was encountered while training the model. Overfitting can be reduced by decreasing model complexity or adding dropout layers. Dropout layers were added in the last few layers to address the problem of overfitting, however, even after adding dropout layers the overfitting problem was not solved.



(a) Dice Loss

(b) Dice Coefficient

Figure 3.24: Dice Values

As we can see from the above figure that the loss function has saturated and yet not reached zero. The Dice coefficient is around 0.8 but it can be improved by adding more convolution layers or reducing the number of levels further and adding more convolution layers. The Dice coefficient value can also be increased by increasing the number of epochs or iterations. This is the loss curve for 30 epochs. However, training has to be stopped when a saturation point is reached otherwise, the graphs start diverging instead of converging which leads to overfitting.

We then tried to use **2.5 D networks**, that are halfway between a 2D network and a 3D network. In a 2.5 D network, a central slice is selected and two neighbouring slices of the central slice are selected so that the network has more knowledge about the images and gets to extract more features from the neighbouring slices. However, the 2.5 D network did not address the overfitting problems, even though it performed slightly better than the 2D network.

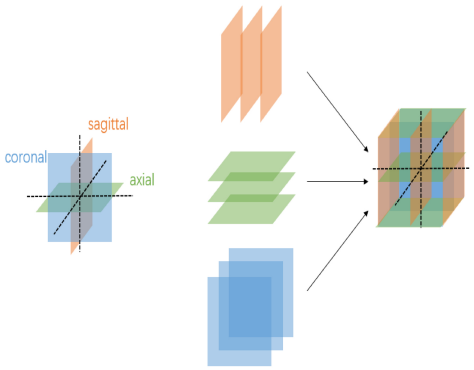
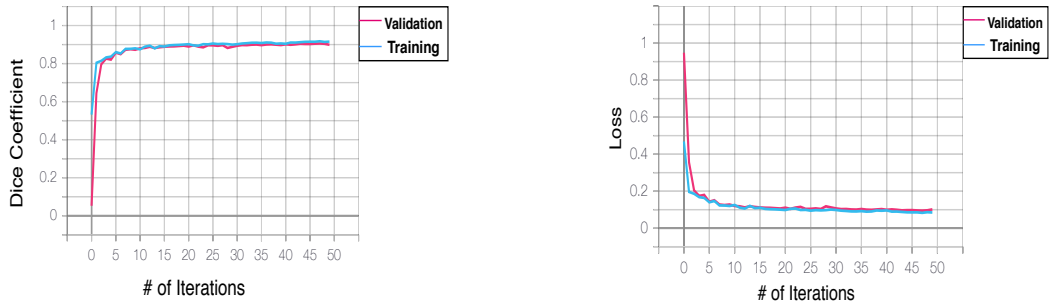


Figure 3.25: 2.5D CNN[40]

At times, the network is biased by the way the input images are processed and can also cause overfitting, the input images need to be shuffled before feeding into the network. The input images were later shuffled before they were fed to the network, this removed the network bias and also solved the overfitting problem. More convolution layers were added to the existing levels to extract more features from the images.

Since the network is 2D, 2D images are fed to the network. We have 1612 training sample points, 180 validation points and 256 test points. The number of epochs is selected as 20 while the batch size is 10. This was first tested for N=2 and then extended to N=6. The Dice coefficient for white matter(0.9), grey matter(0.84) and water(0.87) were high as they constituted about 75% of the total anatomy. Dice coefficients for CSF (0.77), fat(0.8) and eyes(0.3) were low as they constitute the remaining 25% of the anatomy.



(a) Dice Coefficient

(b) Dice Loss

Figure 3.26: Dice Values

Fig 3.20 depicts the Dice values for white matter. As it is clear from both the above images that the over-fitting problem has been solved.

In a study by [21], they achieved Dice coefficients of more than 0.9, this could be due to the fact that they worked on pre-processed 3T datasets and derived the segmentations from T1 and T2 weighted images.

Since, in the current thesis we worked with multi-contrast datasets, different segmentations have been derived from different sequences, which could be a major reason for the discrepancy in Dice coefficient values. To test this hypothesis we trained the network using MP2RAGE images. As explained in chapter-3, the brain segmentations are derived from the MP2RAGE images. We noticed that, if we use MP2RAGE images for training, the Dice coefficient reaches about 0.92 for 50 epochs and about 0.94 for 100 epochs. Furthermore, the MP2RAGE images are automatically intrinsically corrected for image inhomogenities, so it yielded improved segmentations. The difference between the ground truth and predicted images was less as compared with the T1 synergy image. However, we observed that near the borders the predicted image performs better then the ground truth in some of the slices this could be due to the variation in data of the ground truth and potential overlaps between segments from different volunteers. Since, the data we have used is not pre-processed it could also account to the variations along the borders. Sometimes, the predicted images can yield better results than the ground truth.

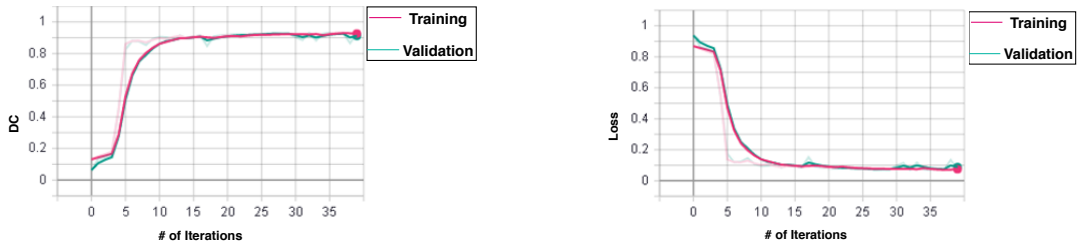


Figure 3.27: Dice Values for MP2RAGE

As we can see in Figure 3.27(a), the Dice coefficient is high as compared to the T1 synergy image in Fig 3.26. The difference between the ground truth image and the predicted image is also less. Hence, if we use the MP2RAGE images for training, we get much better segmentations of the brain as compared to the conventional approach. However, it is not practical to base the training on MP2RAGE images as these images maybe used for research but may not be used for actual clinical purposes. Hence, it is better to use T1 images for training. They are also take less time during acquisition and since, time is of importance here, it is better to use T1 images for training.

The network was trained using T1 images that were bias corrected using N4[35] correction. We observed that the Dice coefficient did not improve much from the T1 synergy images that were not bias corrected. We then tried using in-house B_1 correction and N4 correction, but still there was no improvement in the Dice coefficients, so we conclude that using MP2RAGE images for training have a substantial effect on the Dice coefficients, but these images cannot be used to derive segmentations for SAR predictions. The superior predictions of the network in certain slices could be because there is variability in the ground truth data. For example, we have a CSF label from volunteer 1 and another CSF label from volunteer 2, both these labels may

have certain differences between them. At times it maybe easier for the network to predict one CSF label as opposed to the other one, especially around the borders of the label. So, some slices are easier to segment for the network.

3.4.2. COMPARISON BETWEEN U-NET AND ADAPTED FORKNET

This architecture performs much better than conventional U-net in terms of segmentation. For demonstration, we considered the U-net architecture for N=2,.i.e, white matter and grey matter. For just one tissue it performs as well as the adapted ForkNet, but as we increase the number of tissues the segmentation accuracy of the U-net starts decreasing. To draw an unbiased comparison, the parameters of the network, were the same as the adapted ForkNet. The U-net reported a Dice coefficient of 0.88 for white matter and a Dice coefficient of 0.84 for grey matter. While, for the adapted ForkNet, the Dice coefficient for white matter was 0.9 and that for grey matter is 0.88. Since, the dataset is very small K-fold cross-validation was performed to make sure that the network is robust and not biased towards a particular dataset. Hence, it can be concluded that the adapted ForkNet performs much better than conventional U-net.

K-fold cross-validation: cross-validation is used in machine learning and deep learning to verify that the network is not biased towards a certain set of data points. This is used when the data pool is small, as is in our case. A single parameter called 'k' refers to the number of groups the data is split into. In this thesis we used k as 4. The dataset is shuffled randomly, and the data is split into validation and training sets. For instance, there are four folds, then in the first iteration, the first three folds are training set and the remaining is the test set, in the next set the test data is included in the training set and another set is used for validation, an average dice score is calculated after the four folds have been trained.

Models						
Tissue	White Matter	Grey Matter	CSF	Eyes	Fat	Water
Fold1	0.89	0.89	0.89	0.33	0.87	0.89
Fold2	0.92	0.87	0.87	0.33	0.80	0.87
Fold3	0.89	0.83	0.83	0.32	0.82	0.85
Fold4	0.84	0.84	0.84	0.33	0.83	0.81
Average	**0.89**	**0.85**	**0.85**	**0.33**	**0.83**	**0.85**

Table 3.1: Dice Coefficient for K-fold cross-validation for Adapted ForkNet

S.No.	Tissue	U-Net	Adapted ForkNet
1	White Matter	0.86 ±0.017	0.89 ±0.0193
2	Grey Matter	0.73 ±0.0181	0.85 ±0.0249
3	CSF	0.63 ±0.02	0.77 ±0.0227
4	Eyes	0.25 ±0.0106	0.33±0.0160
5	Fat	0.73 ±0.022	0.83 ±0.0212
6	Water	0.78 ±0.014	0.86 ±0.0316

Table 3.2: Mean Dice Coefficients and Standard Deviation for U-net and Adapted ForkNet

Table 3.2 depicts the differences in Dice coefficients between Adapted ForkNet and U-net proving the superiority of the adapted ForkNet over the conventional U-net. It is clear from the figure that the adapted ForkNet performs much better than the U-net as the Dice coefficient of adapted ForkNet is more than that of U-net when the number of tissues that have to be segmented is increased. The adapted ForkNet has a separate decoder layer for each tissue label which increases the segmentation accuracy.

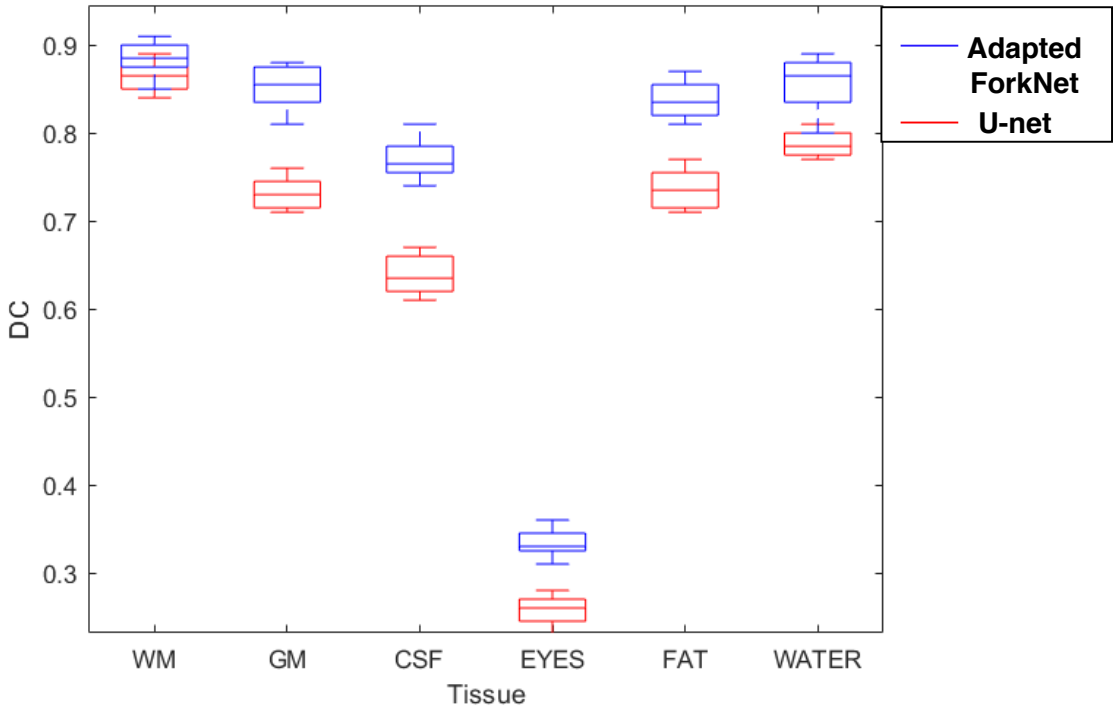


Figure 3.28: Boxplot for U-net and Adapted ForkNet

3.4.3. SAR SIMULATIONS FOR PREDICTED SEGMENTS

It has been established that the adapted ForkNet performs better than the conventional U-net, the aim of the thesis is to predict the SAR values based on the segmentations from the deep learning pipeline. Different decoders are used to segment six tissues hence there can be overlap between a few segments. Since, this is a 2D network, it was only trained for axial direction. It can be further extended to train in the coronal and sagittal planes as well, which is not a part of the thesis. However, axial segmentations provide sufficient segmentations to be used for SAR evaluations. It is then carried out for all the volunteers. The 2D slices are converted to 3D nifti files using a package in Python 3.6. An affine transformation is used to interpolate 2D data to 3D data. Certain overlaps are corrected for using 3D Slicer.[14]. A voxel file is created from the nifti file to be used for SAR simulations in (XFdtd 7.4, Remcom inc., State College, USA). The results are discussed in more detail in chapter-4.

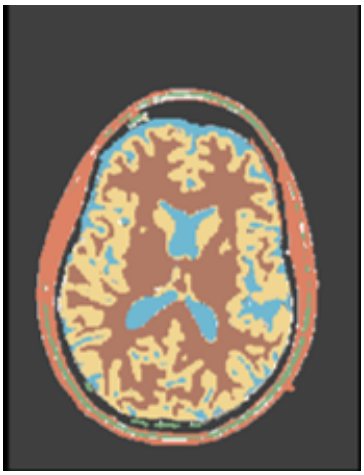


Figure 3.29: Predicted body model overlaid on original body model for 0683

Figure 3.29 depicts a predicted body model overlaid on the original body model for volunteer 4. As we can see from the figure, it is quite clear that the predicted body model is quite similar to the original body model for this volunteer. SAR distributions were computed using the predicted body models and compared with the original body models. The results are discussed in much detail in Chapter-4.

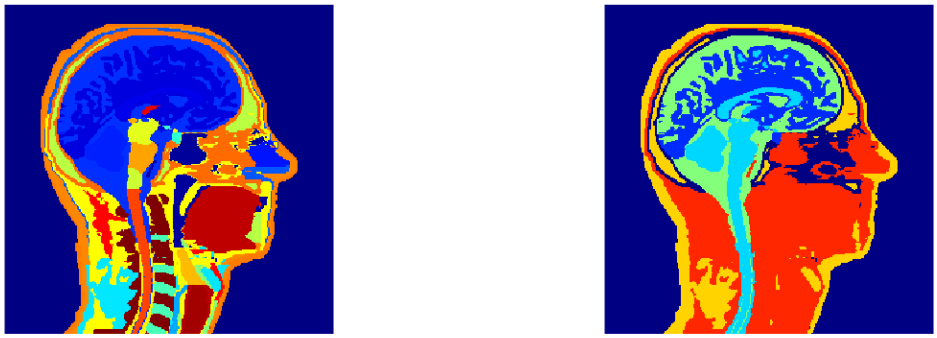
4

RESULTS

This chapter of the thesis describes the results of various findings in this thesis. Section 4.1 illustrates the results of the tissue reduction study. Section 4.2 illustrates the results of the segmentation section. Section 4.3 illustrates the results associated with the creation of numerical body models and SAR evaluations. Lastly, section 4.4 describes deep learning results. Section 4.5 illustrates the comparison between ground truth SAR estimates and predicted SAR estimates.

4.1. DUKE REDUCTION STUDY

In the literature models mentioned in chapter-2, Duke and Ella, the tissues are segmented into 77 different types of the whole body. However, when the full head and shoulders model is considered, the number of segmented tissues reduces to 47. It is a cumbersome process to segment 47 tissues for all volunteers. Hence, it is essential to reduce the number of tissues to a practical number without impairing the corresponding SAR distribution. When patient-specific body models have to be created numerical body model simulations play a pivotal role in determining if the SAR hotspots would be affected by the reduction of tissues. In the current thesis, it was found out that reducing the number of tissues from 47 to 6 does not cause major shift in SAR hotspots. The grouping of these tissues as explained in details in chapter-2 is based on the dielectric properties that these tissues exhibit. Majority of the tissues can be classified as muscle and the remaining tissues are then grouped together on the basis of their dielectric properties.
Subsequently, after confirming that the proposed reduction of tissue classes has only minimal effect on the simulated SAR distribution, we need to segment these tissues to create body models. The segmentation process and deriving of numerical body models from these segmentations id described in details in chapter-3.
The next few results would consist of a comparison of the full model and the reduced model. SAR simulations and B_1^+ maps were generated using this information using (XFdtd 7.4, Remcom inc., State College, USA). A voxel model of the segmentation was converted using 3D slicer by labelling each tissue using a binary label map and then combining them to create a single nifti file which is later created into a vox file using MATLAB.



(a) Detailed Duke containing 47 tissues

(b) Six Tissue Model

Figure 4.1: Comparison of full detailed duke containing 47 tissues and a reduced duke model containing six tissues

In Figure, 4.1(a), we can see that the figures depict the full segmented Duke model with 77 tissues and in fig 4.1(b) the Duke model containing 6 tissues is depicted. These tissues were clustered based on the method explained above and in more detail in chapter-2.

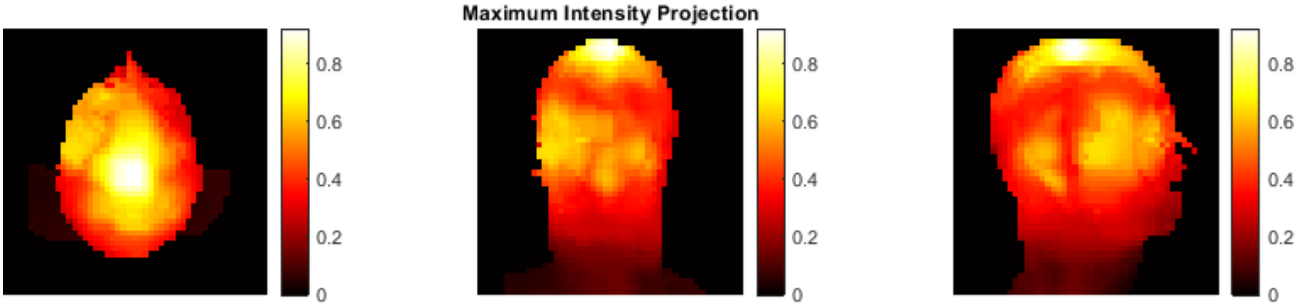


Figure 4.2: 10 g average SAR maps for full duke model

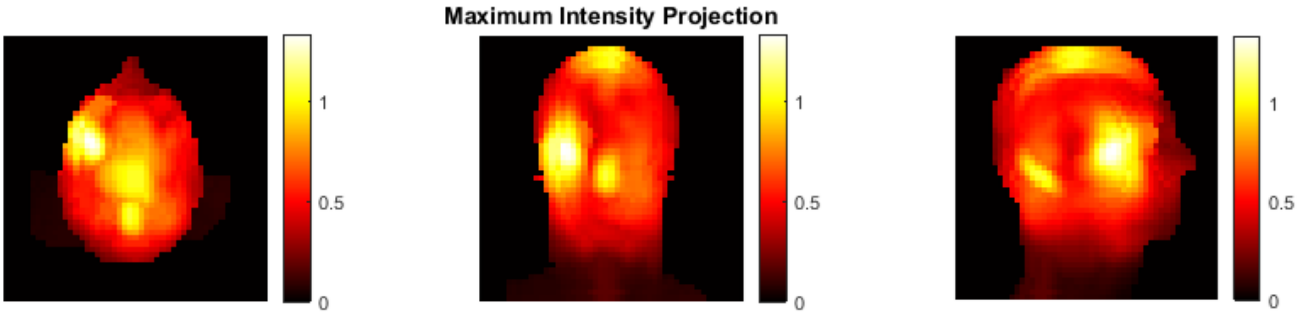
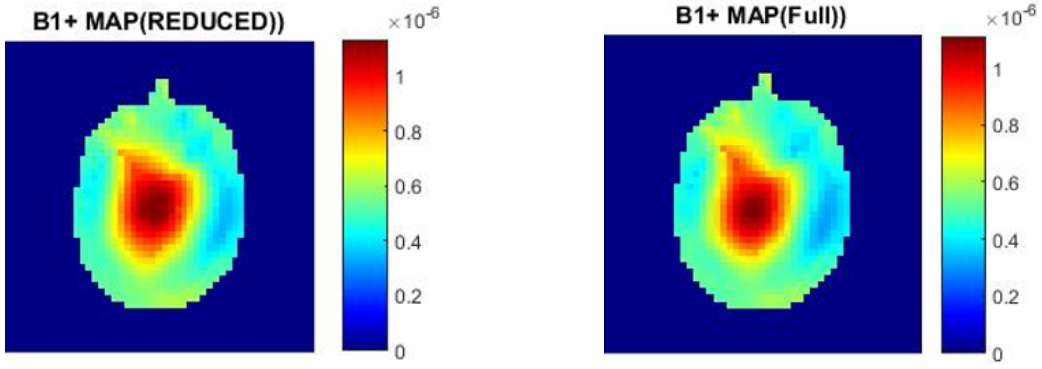


Figure 4.3: 10g average SAR maps for 6 tissue model

Figure 4.2 is 10g averaged SAR of the full duke model and Figure 4.3 10g averaged SAR of the reduced model. The reduced model is similar to the full since various tissues have similar properties and do not contribute much to SAR, so they can be grouped together as one tissue and then segmented.
The SAR simulations were computed using (XFdtd 7.4, Remcom inc., State College, USA). A comparison was drawn between the SAR values of a Duke full model and a reduced model and it was observed that the locations of the hotspots did not change even when the tissues were reduced from 47 to just 6. Suggesting that a six tissue model would result in consistent SAR values as the full model.
The $B1^+$ maps were also computed using (XFdtd 7.4, Remcom inc., State College, USA), and it was observed

that even they strengthen the claim that reducing the tissue model to six tissues does not cause any significant changes in the $B1^+$ field.



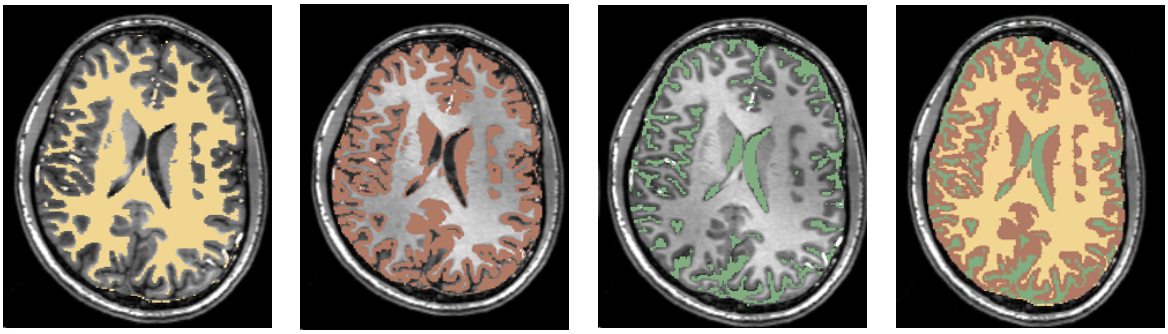
(a) $B1^+$ map(Reduced Model)

(b) $B1^+$ map(Full Model)

Figure 4.4: Comparison of $B1^+$ maps for full duke model and reduced model

4.2. SEGMENTATION

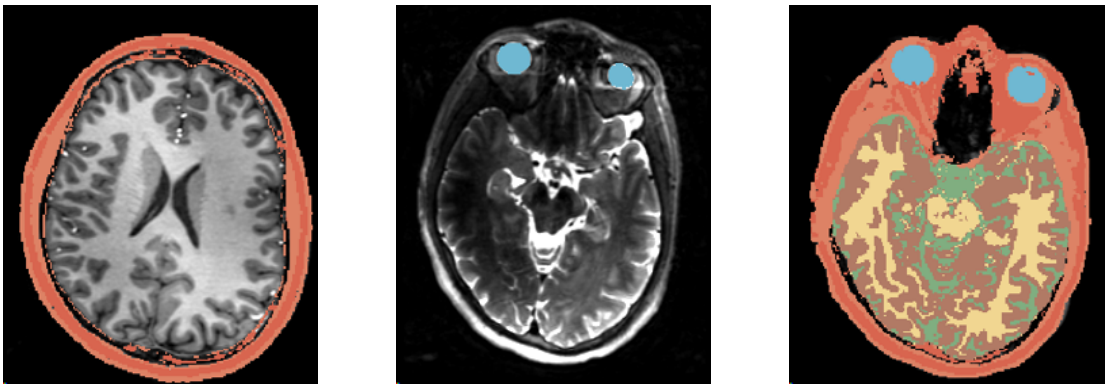
The previous section showed that for creating an accurate numerical body model we need to segment six tissues using the multi-contrast images acquired at 7T MRI scanner. The process of segmentation of these tissues is explained in much detail in chapter-3.



(a) White Matter Segmentation (b) Grey Matter Segmentation (c) CSF Segmentation (d) Full Brain Segmentation

Figure 4.5: Brain Segmentation

In Figure 4.5, the brain is segmented using the MP2RAGE image. CSF, grey matter and white matter are segmented using FSL's FAST toolbox which used k-means clustering algorithm to segment brain tissues. The brain mask has to be manually edited in 3D Slicer before the final brain segmentation. Fig 4.6(d) illustrates the complete brain segmentation and is overlaid on the MP2RAGE UNI image.



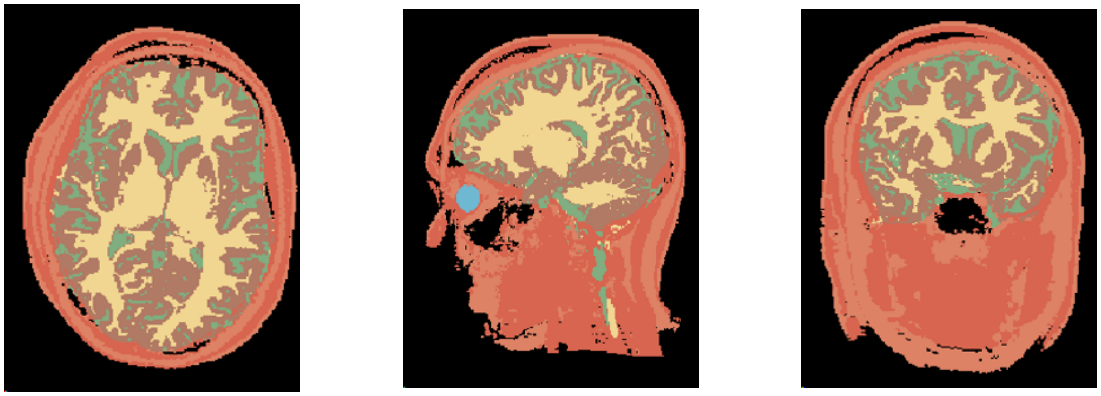
(a) Fat and water Segmentation (b) Eyes Segmentation (c) Complete Segmentation

Figure 4.6: Segmentation of fat and water, eyes and complete segmentation

Figure 4.6(a) illustrates the fat and water segmentations which have been derived from the dixon sequence. Figure 4.6(b) illustrates the eyes segmentation overlaid on a T2 image. 4.6(c) illustrates the final segmentation of six tissues that have been segmented in section 3.2.3. The eyes have been segmented using the T2 image because eyes have a brighter signal on a T2 weighted sequence.

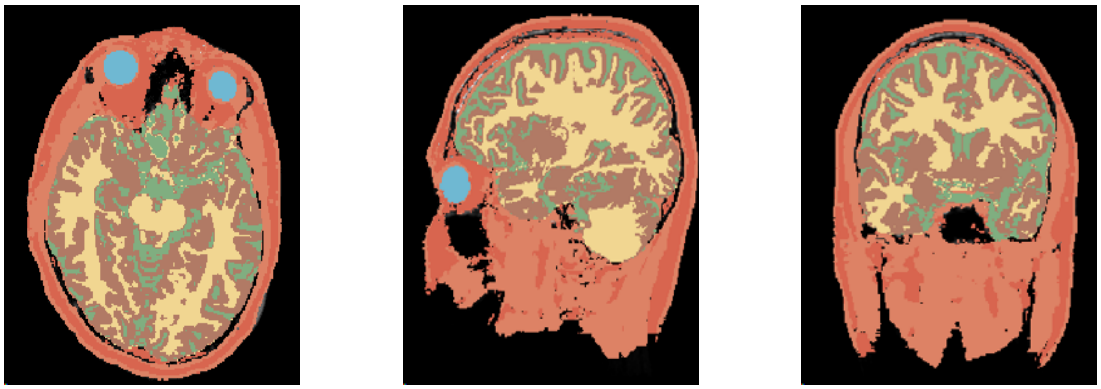
4.3. PERSONALIZED RF BODY MODELS

In this section, the personalized body models of volunteers and their respective B_1^+ maps and SAR estimations are discussed.



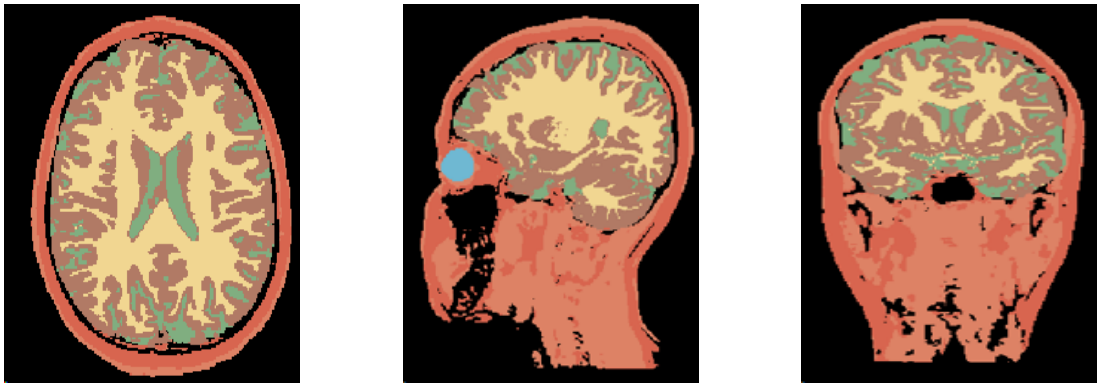
(a) Axial (b) Sagittal (c) Coronal

Figure 4.7: Body Model for Volunteer 1



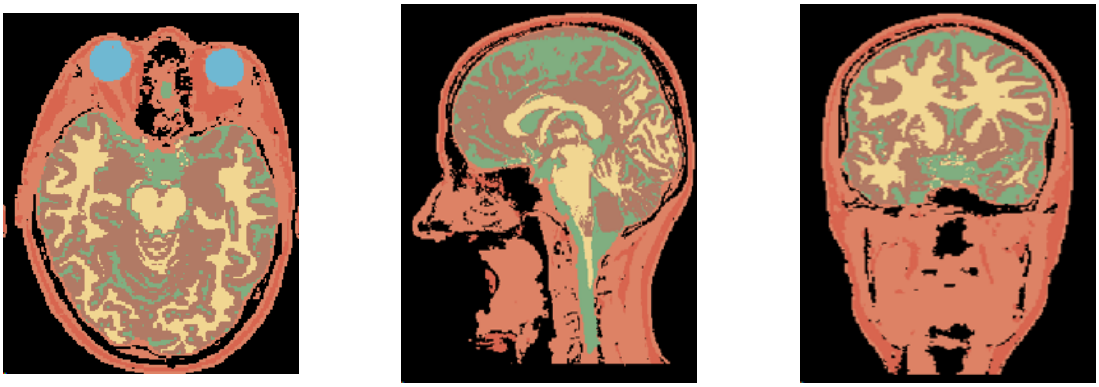
(a) Axial (b) Sagittal (c) Coronal

Figure 4.8: Body Model for Volunteer 2



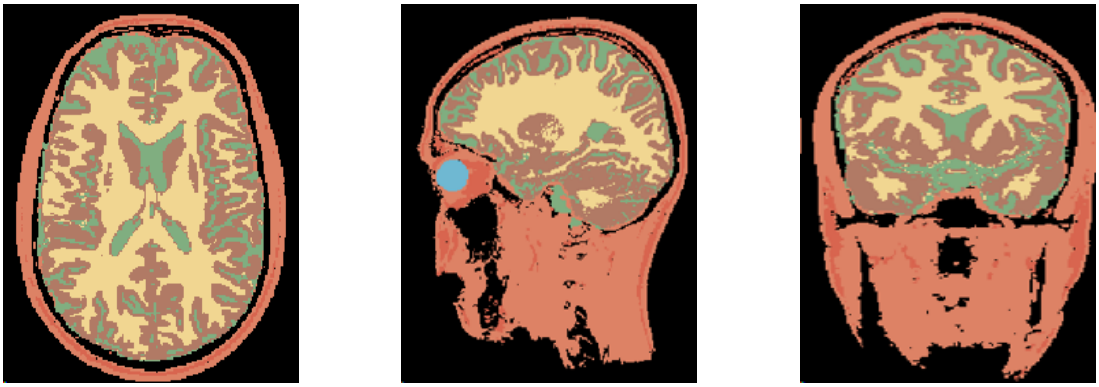
(a) Axial (b) Sagittal (c) Coronal

Figure 4.9: Body Model for Volunteer 3



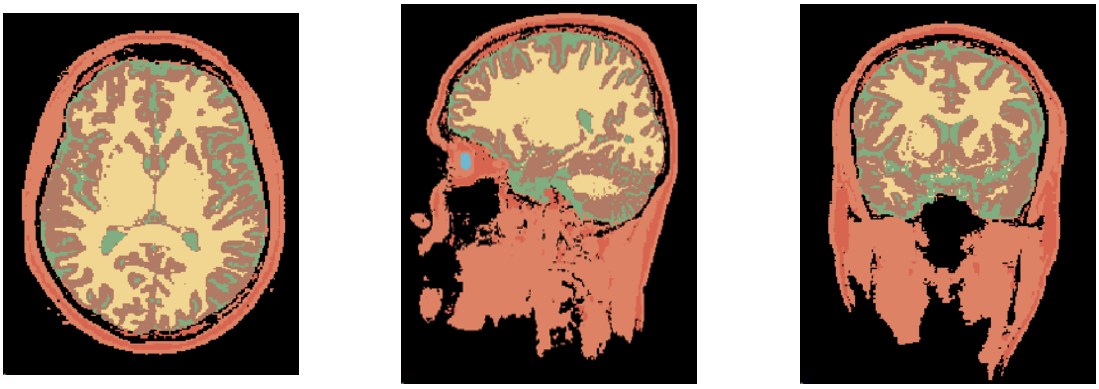
(a) Axial (b) Sagittal (c) Coronal

Figure 4.10: Body Model for Volunteer 4



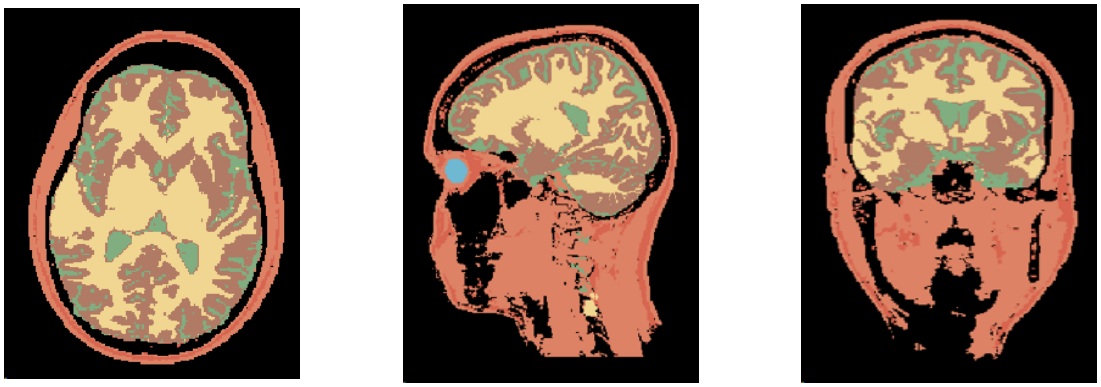
(a) Axial (b) Sagittal (c) Coronal

Figure 4.11: Body Model for Volunteer 5



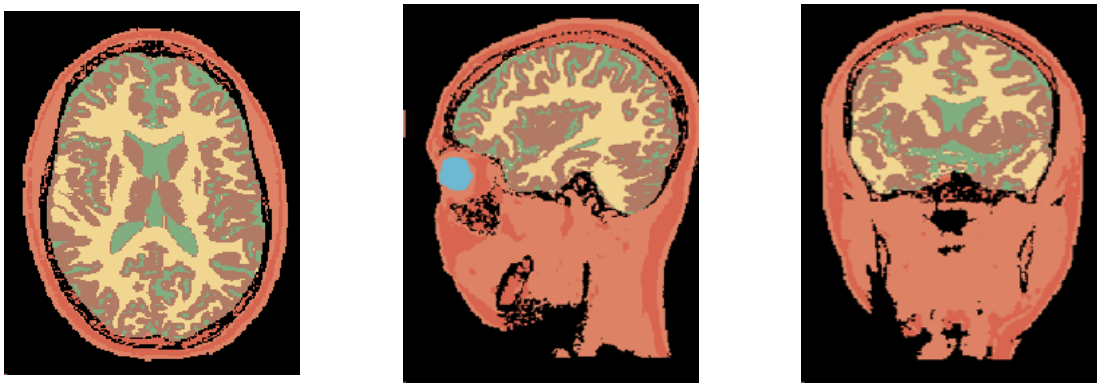
(a) Axial (b) Sagittal (c) Coronal

Figure 4.12: Body Model for Volunteer 6



(a) Axial (b) Sagittal (c) Coronal

Figure 4.13: Body Model for Volunteer 7



(a) Axial (b) Sagittal (c) Coronal

Figure 4.14: Body Model for Volunteer 8

Figures 4.7 to 4.14 depict the body models of various volunteers that are considered in this thesis. The six tissues are segmented as explained in chapter-3. These anatomical models are created from high contrast data which will then be used to compute SAR values. patient-specific RF models are required to accurately estimate SAR values. Since different volunteers have different anatomies and body composition, a personalized body model is created for all the volunteers. The differences in segmentations would lead to differences in SAR hotspots of different volunteers.

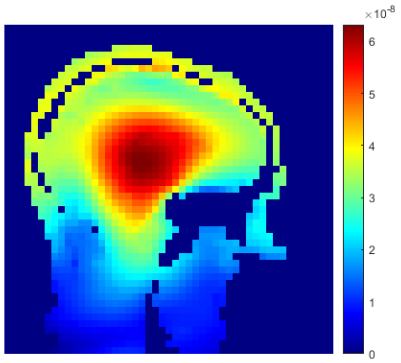


Figure 4.15: Simulated B_1^+ maps for Volunteer 1

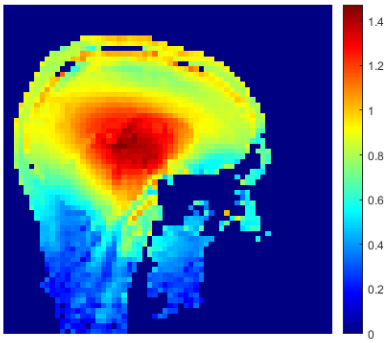


Figure 4.16: Original B_1^+ maps for Volunteer 1

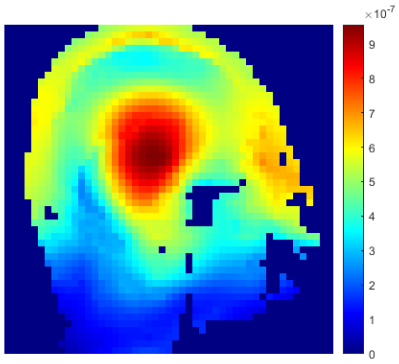


Figure 4.17: Simulated B_1^+ maps for Volunteer 2

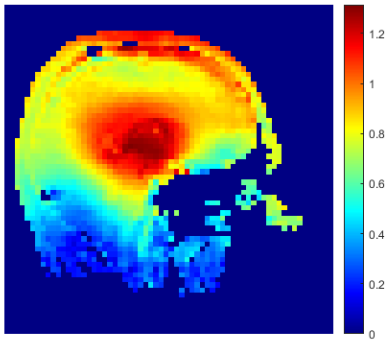


Figure 4.18: Original B_1^+ maps for Volunteer 2

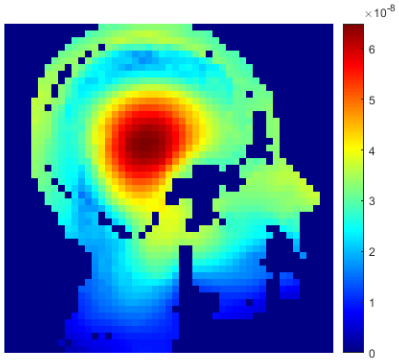


Figure 4.19: Simulated B_1^+ maps for Volunteer 3

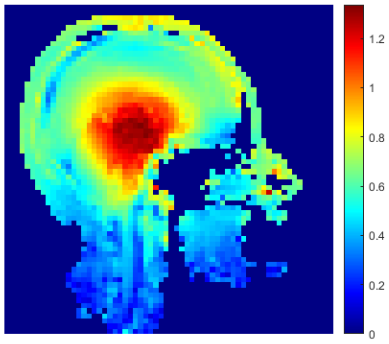


Figure 4.20: Original B_1^+ maps for Volunteer 3

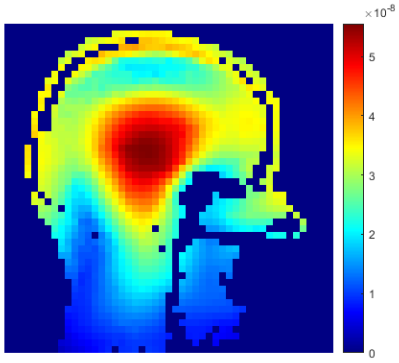


Figure 4.21: Simulated B_1^+ maps for Volunteer 4

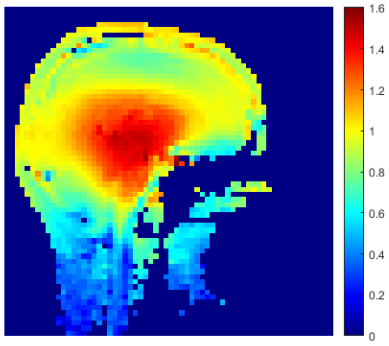


Figure 4.22: Original B_1^+ maps for Volunteer 4

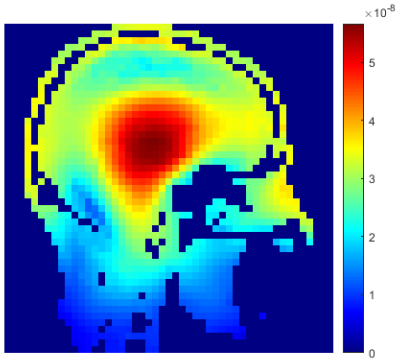


Figure 4.23: Simulated B_1^+ maps for Volunteer 5

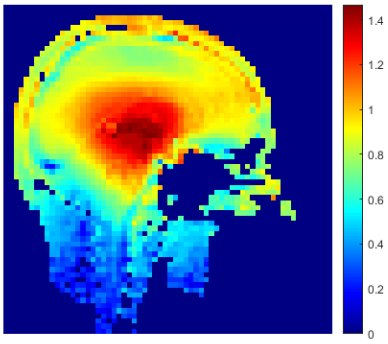


Figure 4.24: Original B_1^+ maps for Volunteer 5

As we can see from Figures 4.15 to 4.24, a comparison has been drawn between B_1^+ maps that have been simulated and B_1^+ maps from the original data. The scale for both the measurements are different from the simulated B_1^+ maps as we use 1W input power, so we just compare the patterns.

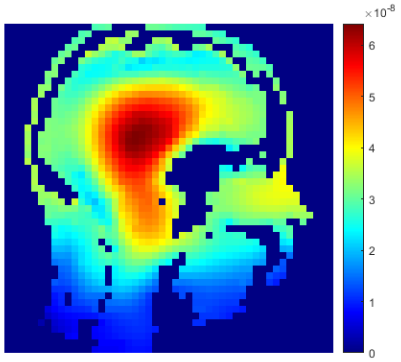


Figure 4.25: Simulated B_1^+ maps for Volunteer 6

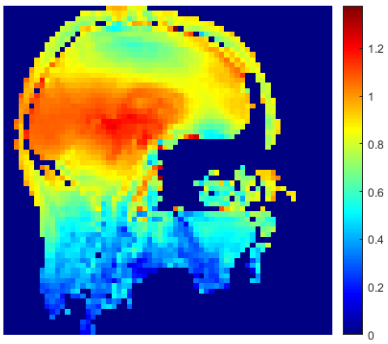


Figure 4.26: Original B_1^+ maps for Volunteer 6

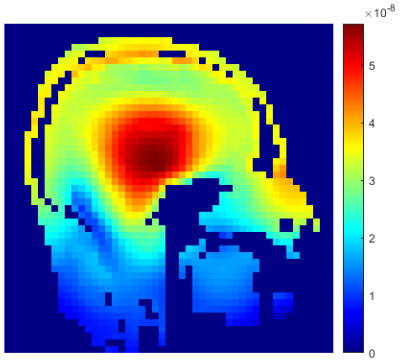


Figure 4.27: Simulated B_1^+ maps for Volunteer 7

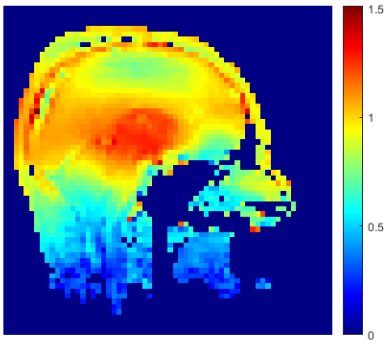


Figure 4.28: Original B_1^+ maps for Volunteer 7

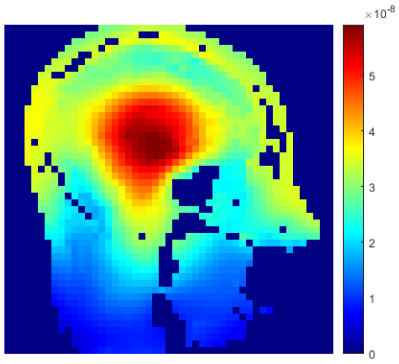


Figure 4.29: Simulated B_1^+ maps for Volunteer 8

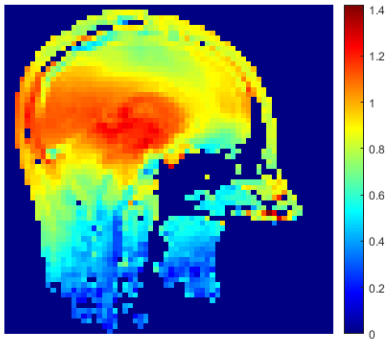


Figure 4.30: Original B_1^+ maps for Volunteer 8

As we can see from Figures 4.25 to 4.30, a comparison has been drawn between B_1^+ maps that have been simulated and B_1^+ maps from the original data. In the above figures, the images are not comparable. B_1^+ maps were not acquired properly for this volunteer, we were not able to acquire the B_1^+ maps with great efficiency.

4.4. DEEP LEARNING

One of the sub-goals of the thesis was to create a deep learning architecture to segment the tissues as described in chapter-3. Six tissues were manually segmented to create numerical body models. The tissues now need to be segmented using a deep learning technique also described in chapter-3. The Adapted ForkNet was used to segment these tissues. ForkNet was used for segmentation. Since it was trained on a 3T dataset there were a lot of problems associated with the 7T images that the network could not handle as it was not equipped to handle the image inhomogenities in 7T. Eventhough, it works well in most cases, it is not as robust for the 7T images as it is for the 3T images. So, a similar network based on the U-net was trained from scratch, keeping in mind the image inhomogeneties in the 7T image. Transfer learning could also have been used instead of training the network from scratch, however, the ForkNet was trained on Mathematica and we did not have enough computational resources to train it using Mathematica.

Figure 4.31: White Matter Segmentation using ForkNet for 3T images

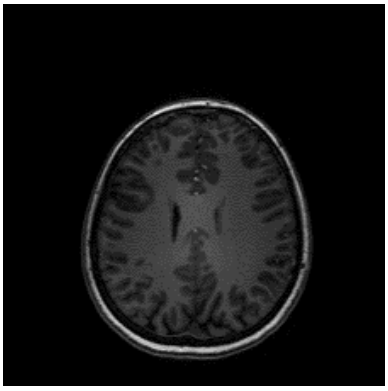


Figure 4.32: Input 7T T1 Synergy Image(B_1 corrected)

Figure 4.33: Grey Matter Segmentation using 7T image

Figure 4.34: Grey Matter segmentation overlaid on the original

Figure 4.35: 7T segmentation

4.4.1. U-NET AND ADAPTED FORKNET

A simple U-net was trained on the MP2RAGE 7T images for one tissue initially. This was performed to check how robust the U-net when introduced to 7T images. Later on, the network would be trained using the T1 and T2 weighted images for multi-class segmentation, i.e, including all six tissues. Some preliminary results of the U-net are discussed in this section.

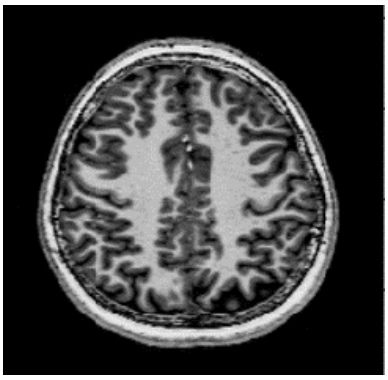


Figure 4.36: Original MP2RAGE UNI image

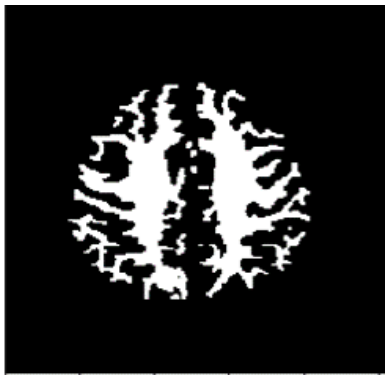
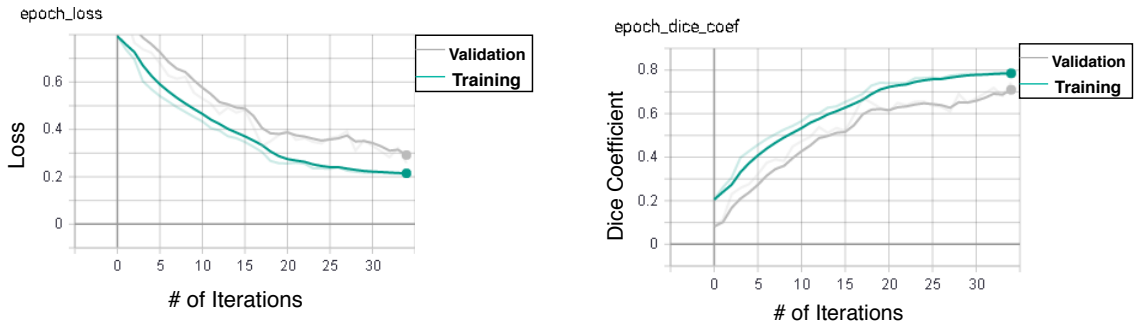


Figure 4.37: White Matter label

Figure 4.38: Predicted image

Figure 4.39: U-net Segmentation

In this thesis, a new network was trained from scratch adapted from [21]. The architecture is similar to the ForkNet as already mentioned in chapter-3. We added additional convolution layers in the ForkNet to extract more features from the image. We call this network Adapted ForkNet. Some of the initial results for the network are illustrated in Figure 4.41



(a) Dice Loss

(b) Dice Coefficient

Figure 4.40: Dice Values

We used dice coefficient as a performance metric for the network. Its value varies from 0 to 1. As we can see from the above image that the dice coefficient almost reached 0.8 for training, while for validation it is about 0.75, which means that training loss is 0.2 and validation loss is 0.25. This is a good value of performance metric since it only varies between 0 and 1. However, the dice coefficient should, in practice, be close to 0.9 for the network to train the minute details around the edges of the images. The training images are T1 weighted images without bias correction so that the network learns to incorporate the bias fields of the images.

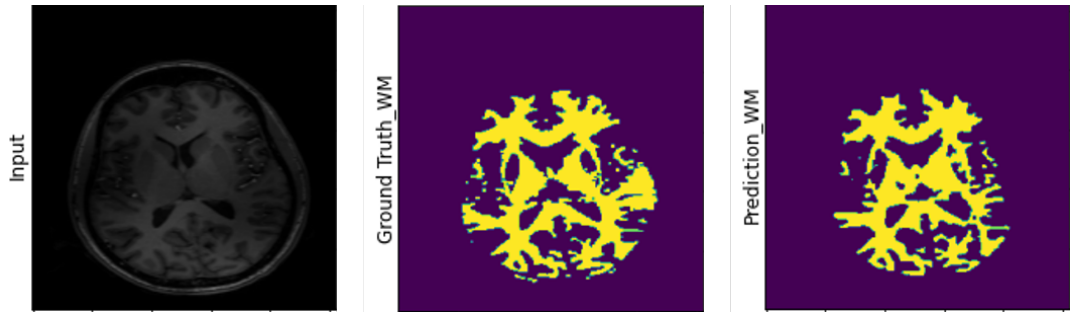


Figure 4.41: Prediction of white matter

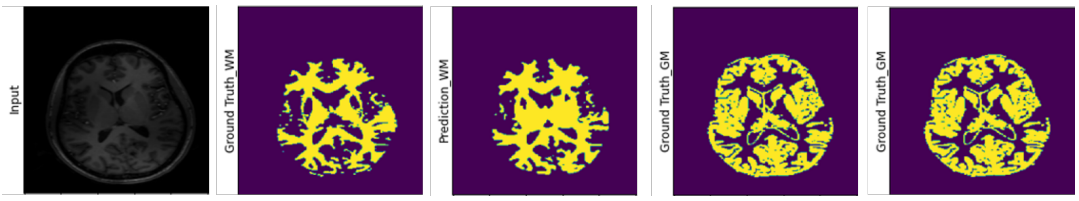
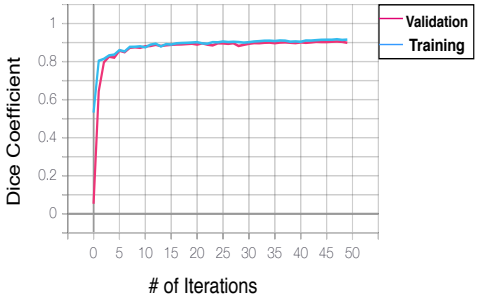
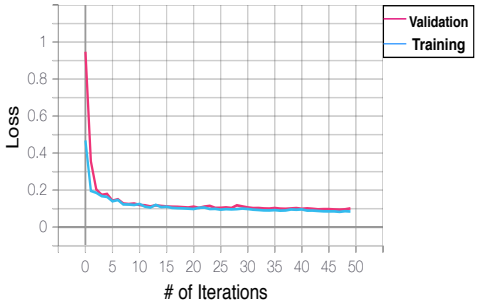


Figure 4.42: Prediction of white matter and grey matter

In Figure 4.42, grey matter is also segmented, the dice coefficient for grey matter was low, hence it is not segmented properly and dilated around the edges.
As described in detail in chapter-3, the overfitting problem of the network was solved using the dropout layers which regularize the network and by shuffling the training data to avoid bias in the network. After the overfitting problem of the network was solved, the network was trained for segmentation of six tissues with each tissue representing a binary label map.



(a) Dice coefficient for White Matter



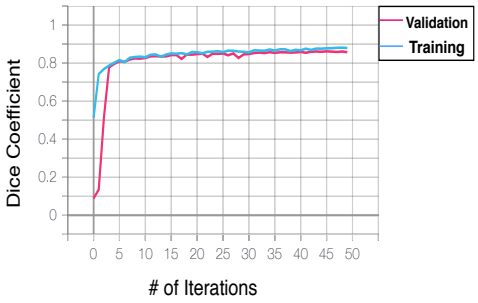
(b) Loss Curve for White Matter

Figure 4.43: Dice Values for White Matter

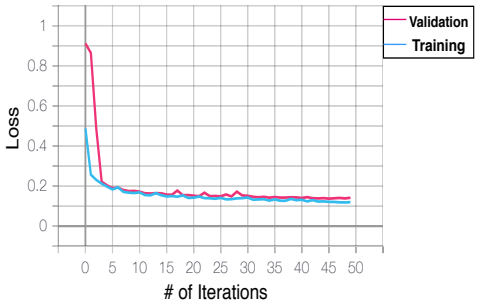
As can be seen from Figure 4.43, the overfitting problem that we had encountered before was addressed. The training and validation curves are so smooth due to the variability of data in the training set. The number of iterations performed was 50 with a batch size of 10. The Dice coefficient is 0.9 for white matter in this case. A large volume of the body's anatomy is represented by white matter, so the Dice coefficient for white matter is high.

Figure 4.44: Comparison between original image, ground truth image and predicted image for White Matter

Figure 4.44, represents the white matter segmentation, the first image is the original T1 image, the second image is the ground truth image derived from high contrast data and the last image is the predicted image that is the output of the network. In certain cases, like this slice, the predicted output segmentation is better than the ground truth, especially near the borders, while in certain cases the predicted output performs worse near the borders of the image, since the Dice coefficient is about 0.9, we cannot expect complete correlation of the predicted output with the ground truth.



(a) Dice coefficient for Grey Matter



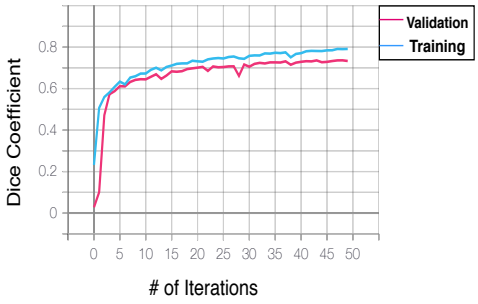
(b) Loss Curve for Grey Matter

Figure 4.45: Dice Values for Grey Matter

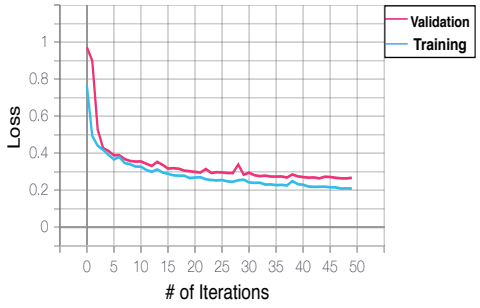
As can be seen from Figure 4.45, the training and validation curves are smooth due to the variability of data in the training set. The number of iterations performed was 50 with a batch size of 10. The Dice coefficient is 0.87 for grey matter in this case. Grey matter comprises of the second largest volume in the anatomy under consideration. So the Dice coefficient for grey matter is also relatively high.

Figure 4.46: Comparison between original image, ground truth image and predicted image for Grey Matter

Figure 4.46, represents the grey matter segmentation, the first image is the original T1 image, the second image is the ground truth image derived from high contrast data and the last image is the predicted image that is the output of the network. In this case, the predicted output is as good as the ground truth image, but in certain cases, the predicted output segmentation is better than the ground truth, especially near the borders, while in certain cases the predicted output performs worse near the borders, since the Dice coefficient is about 0.87, we cannot expect complete correlation of the predicted output with the ground truth. However, the predicted output provides satisfactory results.



(a) Dice coefficient for CSF



(b) Loss Curve for CSF

Figure 4.47: Dice Values for CSF

As can be seen from Figure 4.47, the training and validation curves are smooth due to the variability of data in the training set. The number of iterations performed was 50 with a batch size of 10. The Dice coefficient is 0.77 for CSF in this case. Since, the volume of CSF is less in the anatomy under consideration, which is the head and the neck, the Dice coefficient is slightly lower than that of white matter and grey matter.

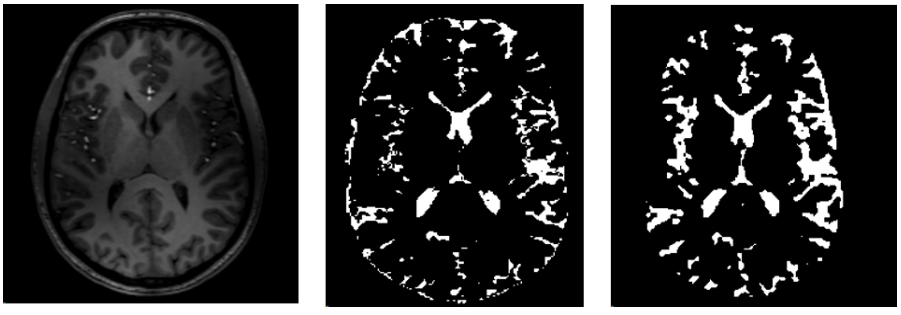
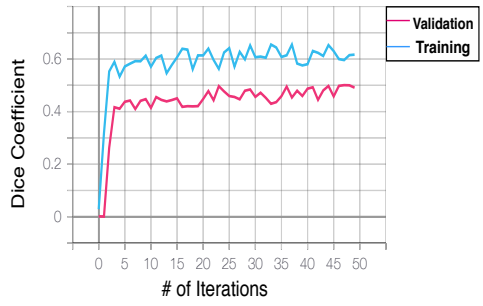
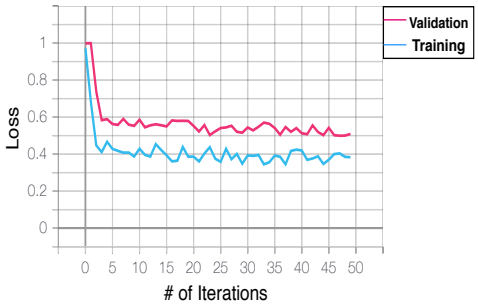


Figure 4.48: Comparison between original image, ground truth image and predicted image for CSF

Figure 4.48, represents the CSF segmentation, the first image is the original T1 image, the second image is the ground truth image derived from high contrast data and the last image is the predicted image that is the output of the network. In certain cases, like the present slice, for instance, the predicted output segmentation is better than the ground truth, especially near the borders, while in certain cases the predicted output performs worse near the borders, since the Dice coefficient is about 0.77, we cannot expect complete correlation of the predicted output with the ground truth. We can expect superior CSF segmentation in this case.



(a) Dice coefficient for Eyes



(b) Loss Curve for Eyes

Figure 4.49: Dice Values for Eyes

As can be seen from Figure 4.49, the training and validation curves are irregular in case of eyes, because eyes represent a very small percentage of the anatomy under consideration, since this region of interest is so small, there cannot be much variability in the data, hence the curves are irregular. The number of iterations performed was 50 with a batch size of 10. The Dice coefficient is 0.33 for eye. Since, the segment sample space is so small, the predicted output is very similar to the ground truth for most cases.

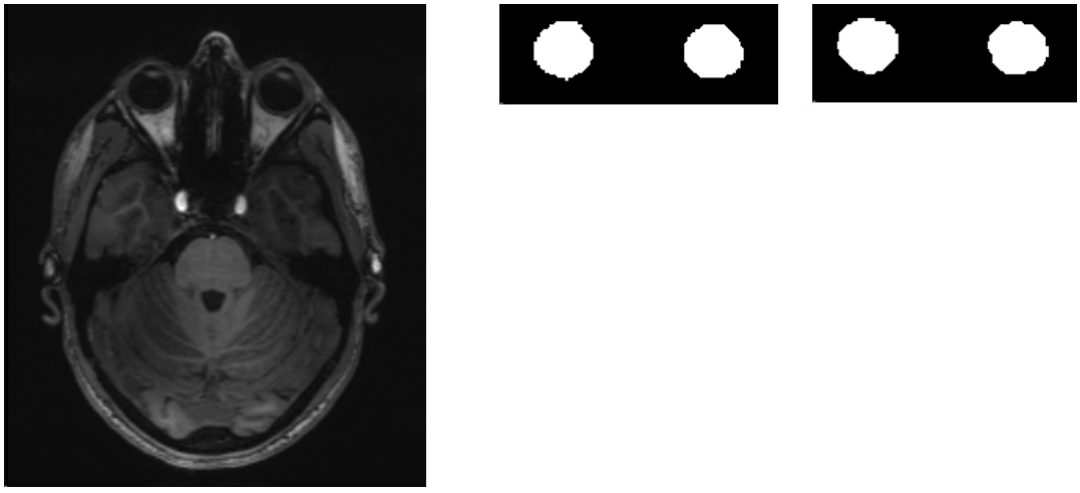
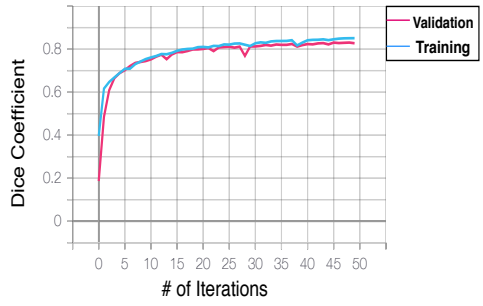
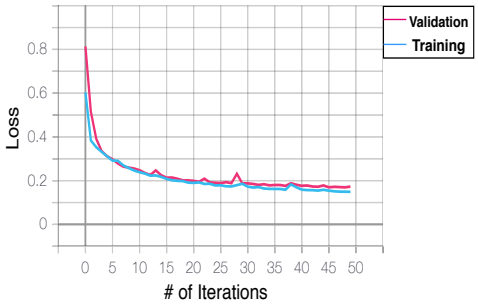


Figure 4.50: Comparison between original image, ground truth image and predicted image for Eyes

Figure 4.50, represents the eyes segmentation, the first image is the original T1 image, the second image is the ground truth image derived from high contrast data and the last image is the predicted image that is the output of the network. As already explained above, since it occupies very little part of the anatomy, the predicted output is very similar to the ground truth in most cases.



(a) Dice coefficient for Fat



(b) Loss Curve for Fat

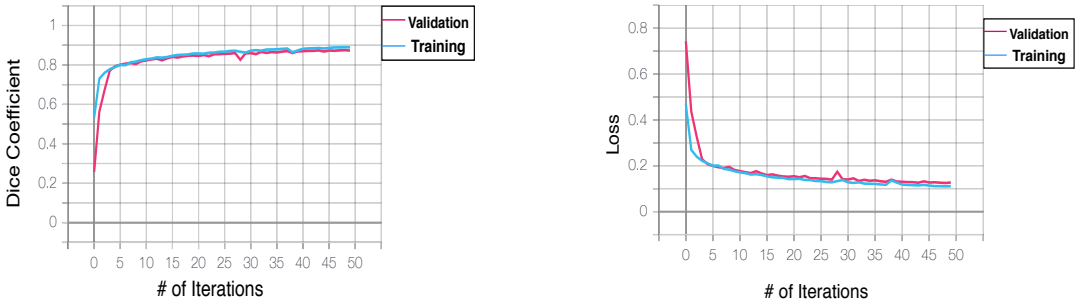
Figure 4.51: Dice Values for Fat

As can be seen from Figure 4.51, the training and validation curves are smooth due to the variability of data

in the training set. The number of iterations performed was 50 with a batch size of 10. The Dice coefficient is 0.84 for fat in this case. Fat occupies a large part of the anatomy under consideration, hence it has a high Dice coefficient as compared to CSF and eyes.

Figure 4.52: Comparison between original image, ground truth image and predicted image for Fat

Figure 4.52, represents the fat segmentation, the first image is the original T1 image, the second image is the ground truth image derived from high contrast data and the last image is the predicted image that is the output of the network. In certain cases, like the present slice, for instance, the predicted output segmentation is the same as the ground truth, while in certain cases the predicted output may perform worse or better near the borders.



(a) Dice coefficient for Water(Muscle)

(b) Loss Curve for Water(Muscle)

Figure 4.53: Dice Values for Water

As can be seen from Figure 4.53, the training and validation curves are smooth due to the variability of data in the training set. The number of iterations performed were 50 with a batch size of 10. The Dice coefficient is 0.89 for water in this case. Water occupies a large part of the anatomy under consideration, hence it has a high Dice coefficient as compared to CSF, fat and eyes.

Figure 4.54: Comparison between original image, ground truth image and predicted image for Water

Figure 4.54 represents the water segmentation, the first image is the original T1 image, the second image is the ground truth image derived from high contrast data and the last image is the predicted image that is the output of the network. In certain cases, like the present slice, for instance, the predicted output segmentation

[performs worse than the ground truth, while in certain cases the predicted output may perform better or almost as good as the ground truth near the borders.

4.5. COMPARISON BETWEEN SAR VALUES OF GROUND TRUTH AND PREDICTED SAR VALUES

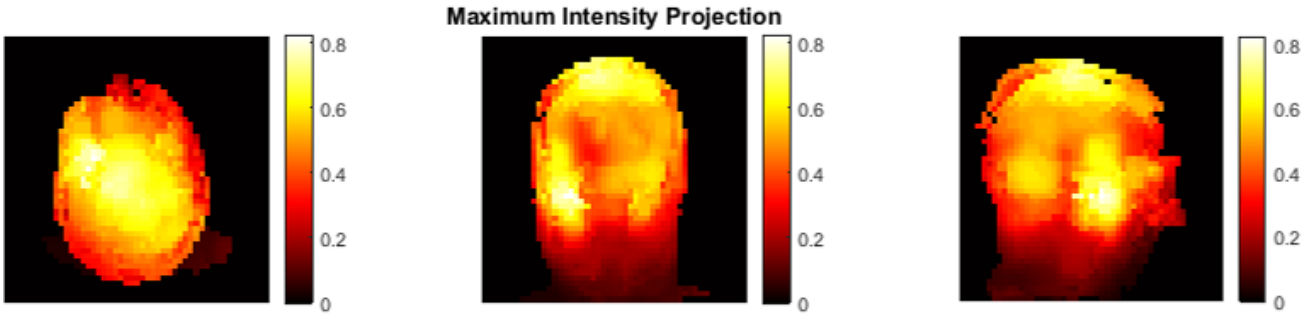


Figure 4.55: 10g Averaged Original SAR Map for Volunteer 1

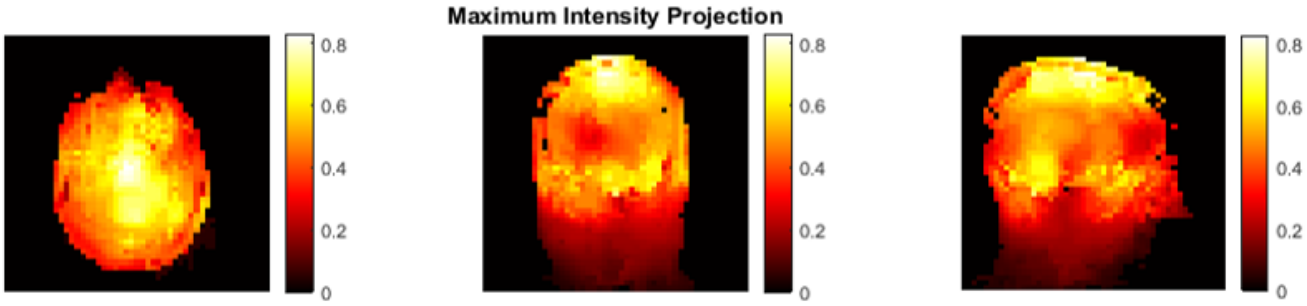


Figure 4.56: 10g Averaged Predicted SAR Map forVolunteer 1

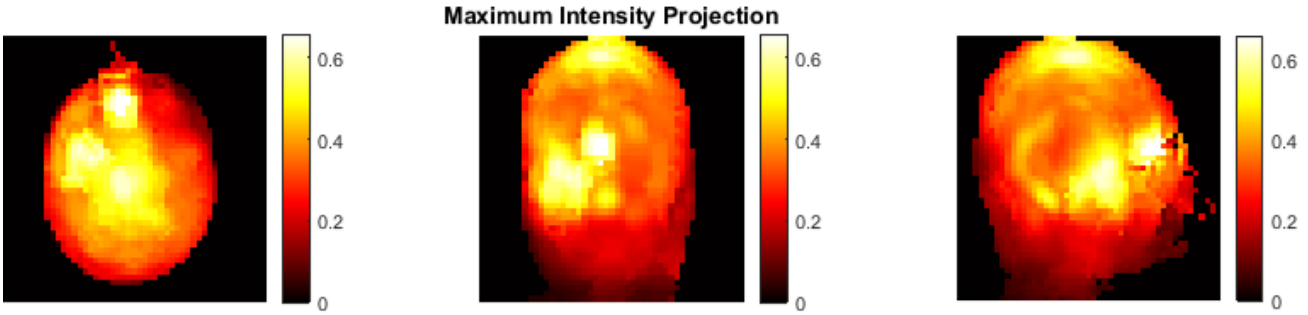


Figure 4.57: 10g Averaged Original SAR Map for Volunteer 2

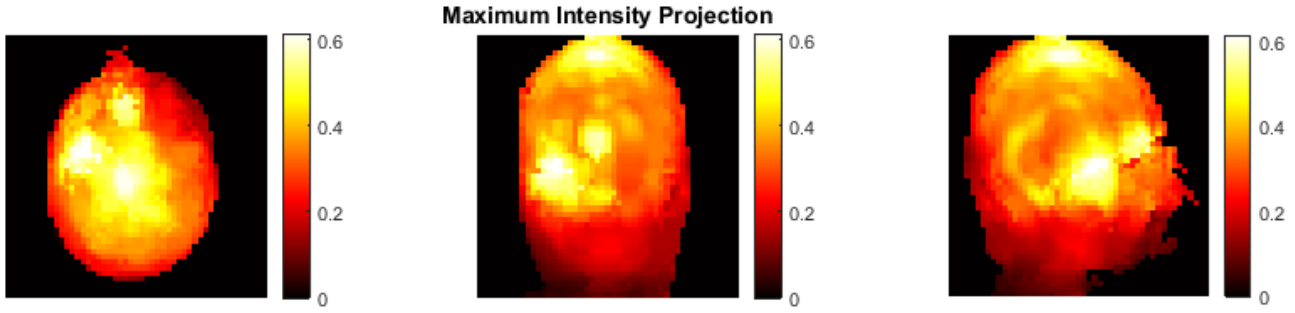


Figure 4.58: 10g Averaged Predicted SAR Map for Volunteer 2

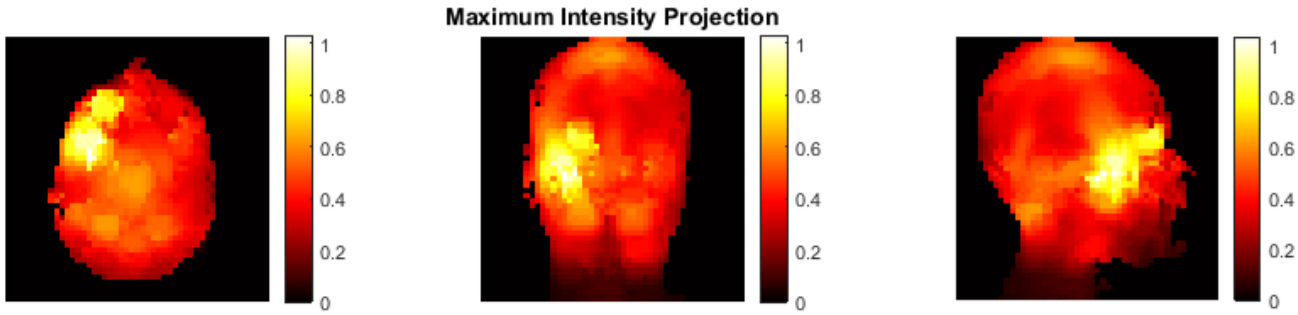


Figure 4.59: 10g Averaged Original SAR Map for Volunteer 3

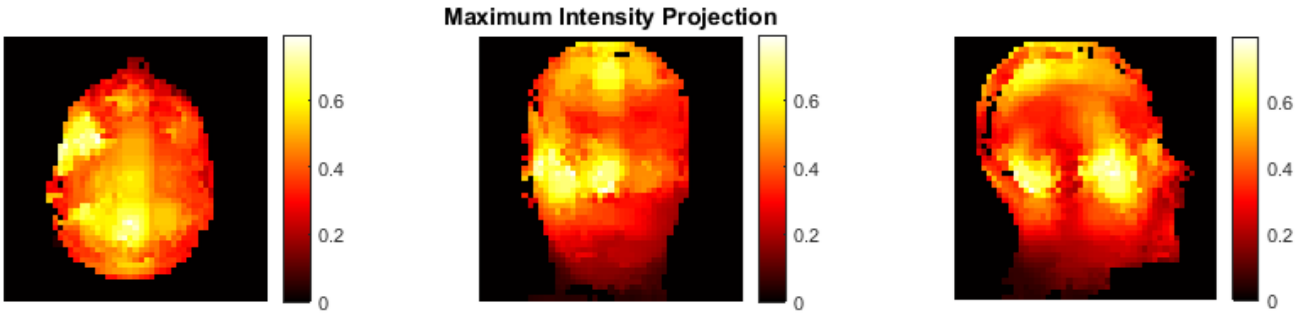


Figure 4.60: 10g Averaged Predicted SAR Map for Volunteer 3

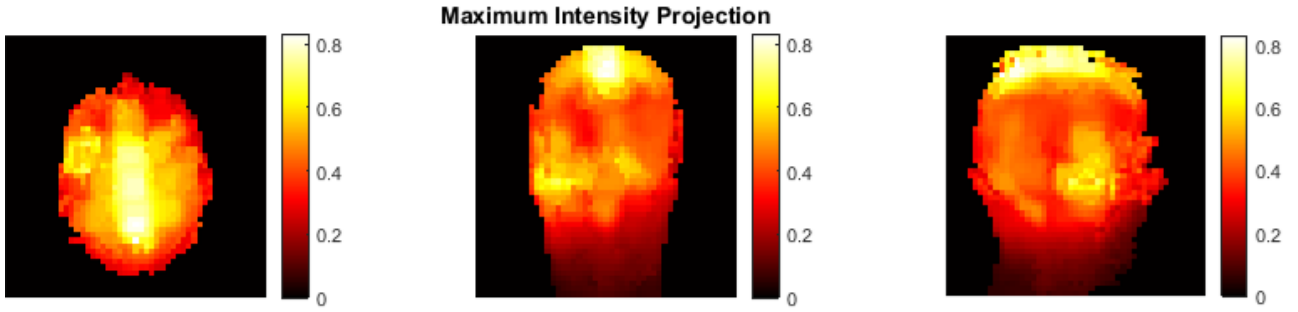


Figure 4.61: 10g Averaged Original SAR Map for Volunteer 4

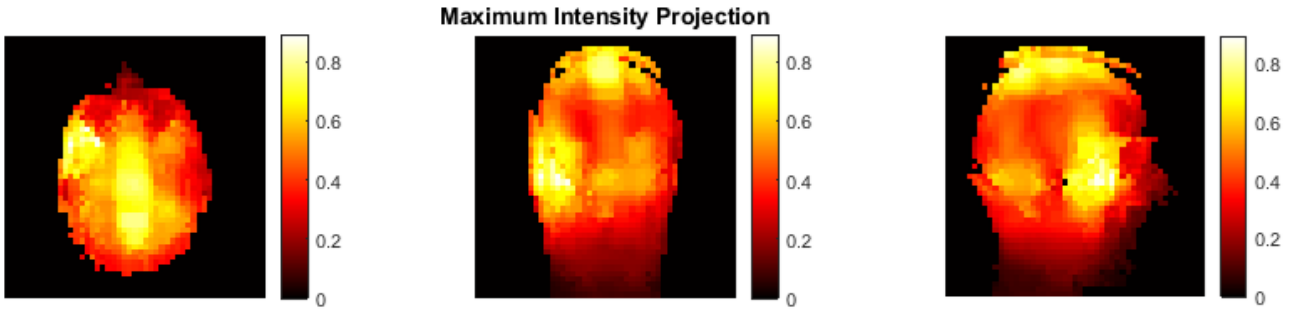


Figure 4.62: 10g Averaged Predicted SAR Map for Volunteer 4

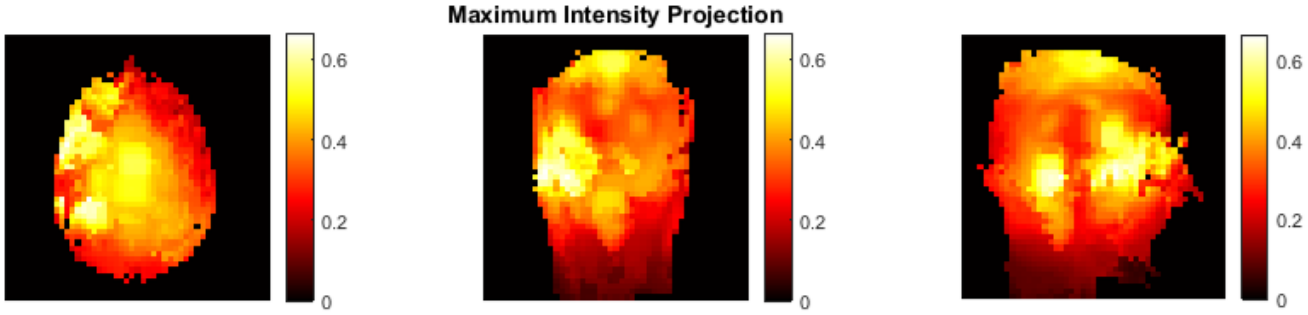


Figure 4.63: 10g Averaged Original SAR Map for Volunteer 5

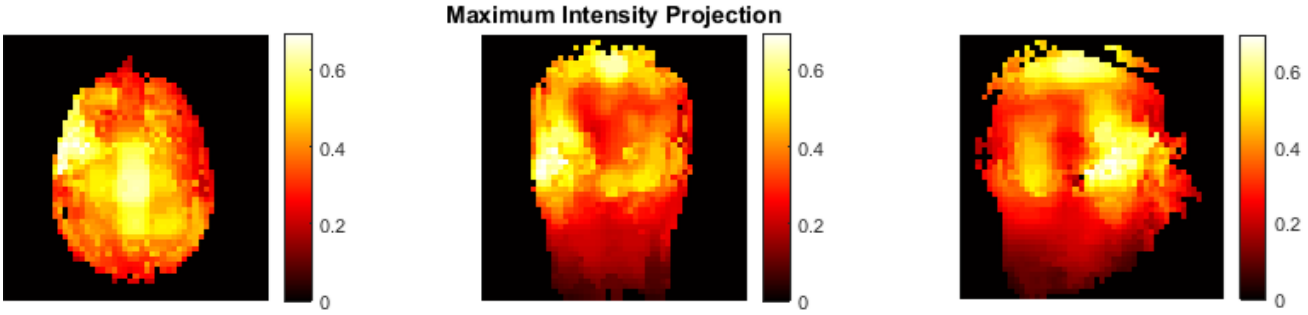


Figure 4.64: 10g Averaged Predicted SAR Map for Volunteer 5

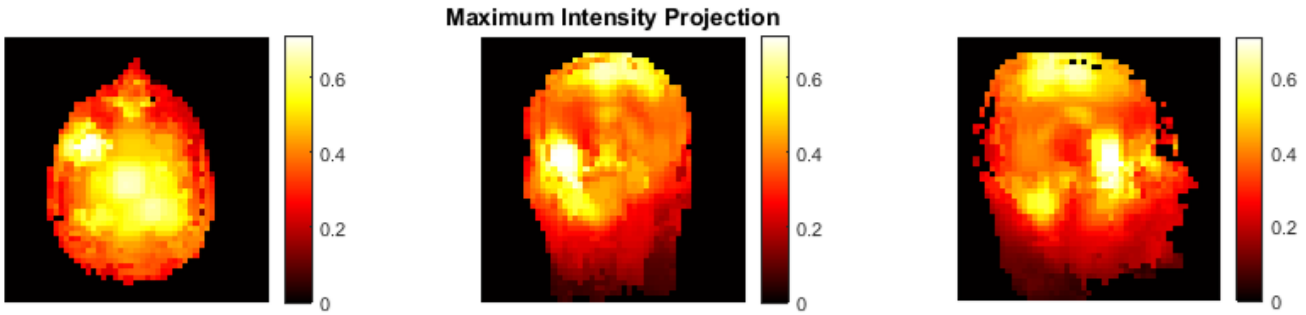


Figure 4.65: 10g Averaged Original SAR Map for Volunteer 6

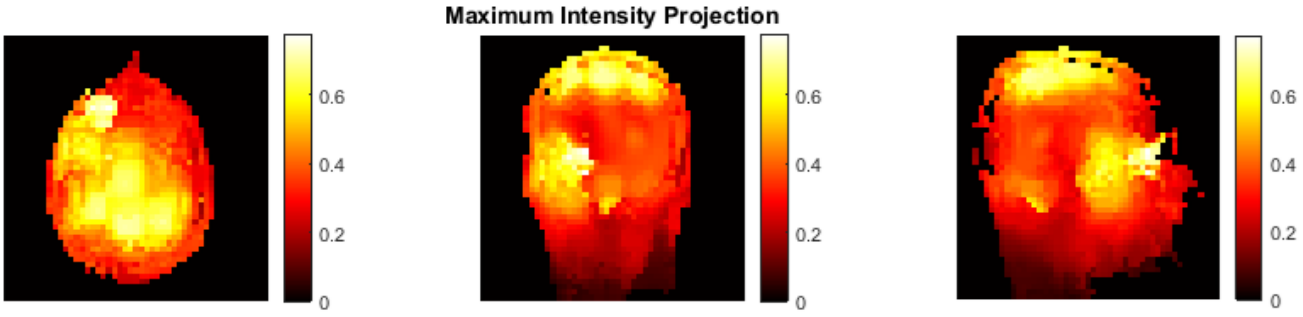


Figure 4.66: 10g Averaged Predicted SAR Map for Volunteer 6

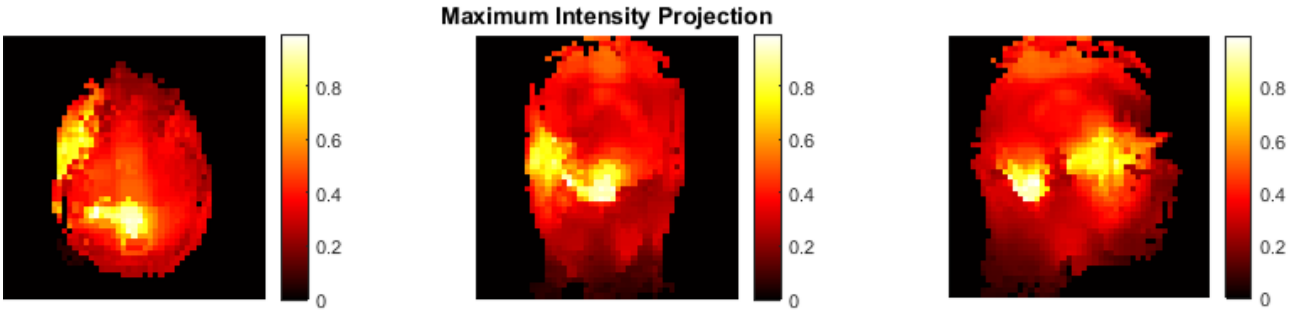


Figure 4.67: 10g Averaged Original SAR Map for Volunteer 7

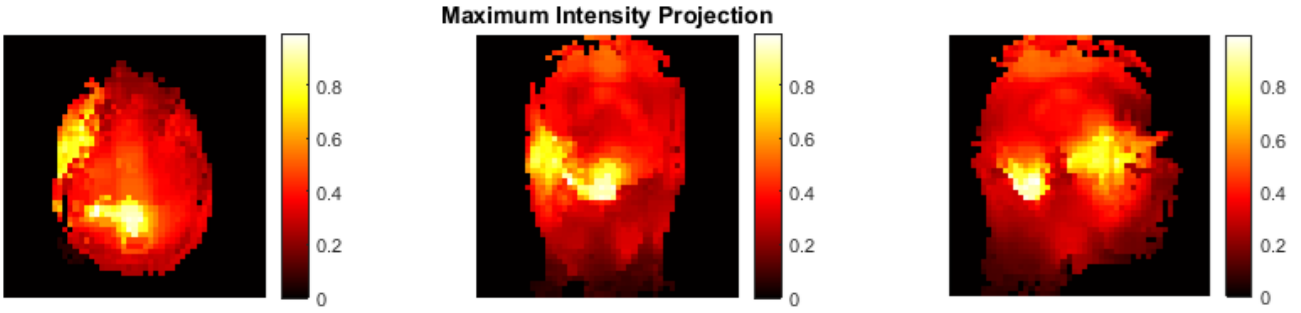


Figure 4.68: 10g Averaged Predicted SAR Map for Volunteer 7

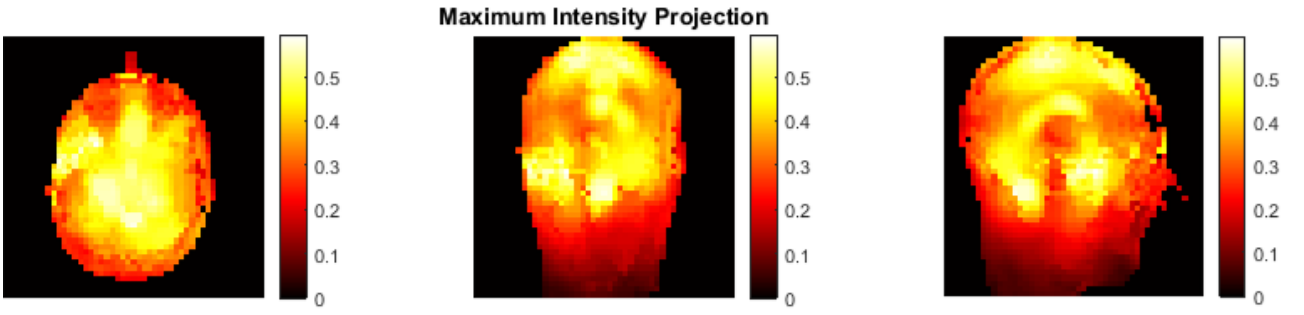


Figure 4.69: 10g Averaged Original SAR Map for Volunteer 8

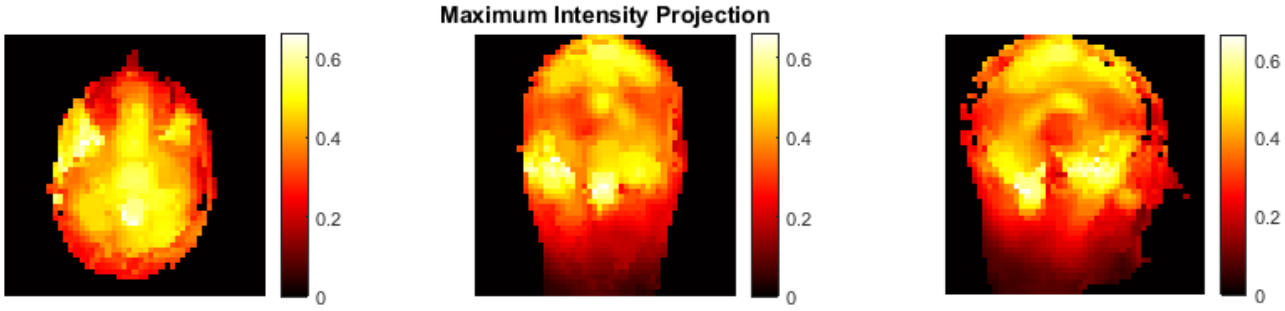
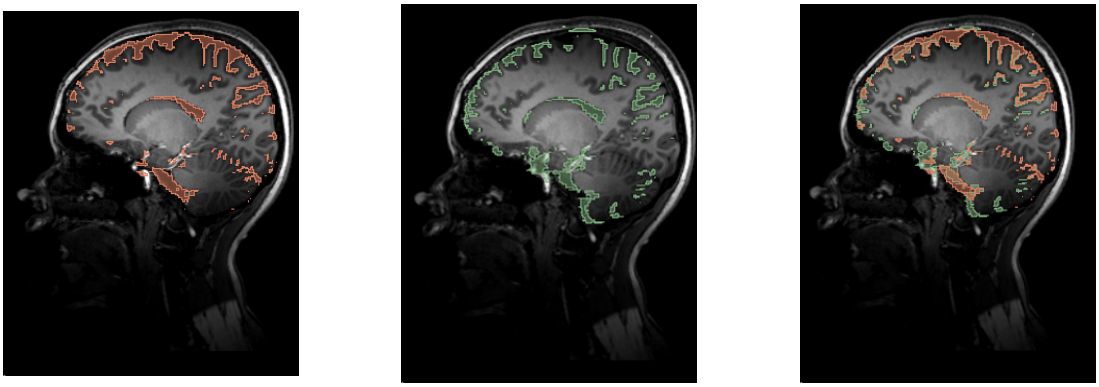


Figure 4.70: 10g Averaged Predicted SAR Map for Volunteer 8

4.5.1. DISCUSSION OF COMPARISON OF SAR DISTRIBUTIONS

After analysing the data from Figures 4.55 to Figure 4.70, it can be interpreted that the SAR hotspots obtained from the predicted body models are very similar to the SAR hotpots in the original body model. Even though there are a few differences between the ground truth and the predicted segmentations they do not influence the SAR values by a great extent. The segmentations are performed using a 1mm resolution image, while the SAR values are evaluated using a 2x2x2mm voxelised grid. Insignificant changes in segmentations at 1mm level would not affect SAR on a 5mm level. However, if there are changes in tissues that have high conducitvity, for example, CSF, the SAR hotspots seem to change. To illustrate this, let us look at the SAR hotspots of the volunteer 3, the network was not able to segment CSF properly for this volunteer, maybe because it was a difficult test set for the network to predict. The CSF segmentation for this volunteer was not satisfactory which led to changes in SAR hotspots.



(a) CSF Ground Truth Segmentation

(b) CSF Predicted Segmentation

(c) CSF segmentation overlaid

Figure 4.71: CSF Segmentation for Volunteer 3

Figure 4.71 illustrates that the CSF segmentation in volunteer 3 was not predicted correctly which led to incorrect SAR estimates for this particular volunteer. The personalised body models and SAR maps are depicted here for eight volunteers. There is no reference to compare these SAR values with as we do not know the actual SAR value. The SAR patterns are similar to the reduced Duke model, but also different for different volunteers. The basic SAR pattern should not, deviate much from the Duke model. However, it is dependent on the anatomy and body composition of different volunteers. Hence, comparing it with the Duke model would not be the right approach. Small errors in the segmentation process, however, do not influence SAR values. Since the ground truth segmentations are done using 1mm images and the voxelised grid of the SAR computations is 2mm, small errors in segmentation at 1mm model do not propagate into the 2mm voxelised grids. However, if the segmentation at the 1mm model is not robust or inaccurate, the errors propagate into the 2mm model and lead to incorrect SAR estimates. Hence, it remains essential to make the ground truth as accurate as possible.

5

Discussion

5.0.1. Discussion and Conclusion

UHF MRI scanners are used for clinical research to image minute structures of the human body. However, the heat dissipated in the body due to these high RF field remains a major bottleneck in the performance of the UHF MRI systems. The goal of the thesis was to accurately predict SAR distributions from images acquired at the 7T scanner and then compare the SAR distributions using a deep learning algorithm. To calculate SAR distributions, we need to create numerical body models from the acquired data. These body models must be patient-specific to consider the variability in population. To create these body models, we need to determine the number of tissues that must be segmented. A comparative study was conducted to determine the number of tissues that must be segmented so that the SAR distributions are not affected. We concluded that the number of tissues that have to be segmented is six. This is a trade-off between SAR accuracy and ease of segmentation.

Manual and automatic segmentation algorithms must be used to create a consistent ground truth model from multi-contrast data. It was concluded in 3.3.3 that MP2RAGE sequence provided the best segmentation for brain tissues (grey matter, white matter and CSF), T2 weighted image was later used to manually correct CSF. Dixon sequence was used to segment fat and water(muscle) and T2 weighted image was used to segment eyes. It was also concluded that a 1mm sequence would be used for segmentation. Numerical body models were computed and SAR distributions for the ground truth model were computed. It is essential to use deep learning algorithms for segmentation to create numerical body models.

Deep Learning algorithms were used to automatically segment these tissues. A comparison was drawn between conventional U-net and adapted ForkNet and it was evident that the adapted ForkNet used in this thesis was superior compared to the conventional U-net. Hence, we concluded in 3.4.2 that the adapted ForkNet was to be used as the final network architecture.

The SAR values for the predicted model were computed using (XFdtd 7.4, Remcom inc., State College, USA) and then compared with the original SAR values. It was concluded that the predicted SAR hotspots are comparable to the original SAR hotspots despite of errors in segmentation. Since, the SAR values are computed using a 2x2x2 mm voxel grid, small errors in the 1mm model can be ignored and do not influence final SAR values. Hence, the SAR predictions using the body models from deep learning algorithms are like the SAR predictions using the original numerical body model.

5.0.2. Limitations of the Thesis

- **Data:** We use the images acquired from 8 volunteers. Increasing the sample space would give us more sample points and more data for the network to train on. We used T1 1mm images for training the network. T1 images take about 3 minutes in acquisition, a localizer image would be quick to acquire and can be used in place of the T1 1mm image.

- **Segmentation:** There could be potential errors in the ground truth segmentation. The errors could also propagate into the deep learning network, hence reducing the efficiency of the network. For instance, the flip angle towards the base of the neck was not consistent and led to the variation in fat and water segmentation. In some volunteers, we could only image a portion of the neck.

- **Gold Standard for SAR values:** We are predicting the SAR values for the ground truth model based on the assumption followed from the duke reduction study that a reduced body model does not affect SAR to a great extent. However, there is uncertainty related to the true SAR values which are still unknown. We can only B1+ maps for simulated and original data.

5.0.3. FUTURE WORK

There are certain limitations associated with the thesis, we can improve these limitations in the future.

- **Data:** Acquiring more data would already improve the accuracy of the segmentation as the network would have more data points to learn from. Acquisition of localizer MRI images would seamlessly integrate the algorithm with MRI scanners for correct prediction of SAR values
- **Segmentation:** As discussed earlier, segmentation errors arise due to image inhomogeneities in 7T images, so it is important to correct the bias field. Better algorithms can be developed for bias field correction to address this problem. Acquisition protocol can be improved to have a larger signal in the neck area which would improve the fat and water segmentation.
- **Deep Learning:** The network was only trained for the axial direction, training the network in the other two planes and combining the results should give better predictions. Multi- contrast dataset may also improve predictions of the network. For example, we could use both T1 and T2 images for training which could improve the segmentation for CSF, for instance. We could also use a multipath 2.5 D CNN network architecture, which would train the network using the slices from all the planes using different encoders and then combining the features so that there is only one input for the decoder. We could also experiment with the loss function of the network. A combination of two loss functions could give better segmentation accuracy, especially near the borders of the image.

Despite the limitations associated with the thesis we could conclude that the SAR distributions of the predicted numerical body model were comparable with the SAR distributions of the ground truth numerical body model. With improved algorithms in the field of bias correction of 7T images we can hope that the segmentation would improve also improving the accuracy of the network.

BIBLIOGRAPHY

[1] D. Bhattacharyya and T.-h. Kim, *Brain tumor detection using mri image analysis,* in *International Conference on Ubiquitous Computing and Multimedia Applications* (Springer, 2011) pp. 307–314.

[2] J. M. Ahn and G. Y. El-Khoury, *Role of magnetic resonance imaging in musculoskeletal trauma,* Topics in Magnetic Resonance Imaging **18**, 155 (2007).

[3] S. Trattnig, E. Springer, W. Bogner, G. Hangel, B. Strasser, B. Dymerska, P. L. Cardoso, and S. D. Robinson, *Key clinical benefits of neuroimaging at 7 t,* Neuroimage **168**, 477 (2018).

[4] C. M. Deniz, *Parallel transmission for ultrahigh field mri,* Topics in magnetic resonance imaging: TMRI **28**, 159 (2019).

[5] P. A. Bottomley and W. A. Edelstein, *Power deposition in whole-body nmr imaging,* Medical physics **8**, 510 (1981).

[6] A. Christ, W. Kainz, E. G. Hahn, K. Honegger, M. Zefferer, E. Neufeld, W. Rascher, R. Janka, W. Bautz, J. Chen, *et al.*, *The virtual family—development of surface-based anatomical models of two adults and two children for dosimetric simulations,* Physics in Medicine & Biology **55**, N23 (2009).

[7] S. Wolf, D. Diehl, M. Gebhardt, J. Mallow, and O. Speck, *Sar simulations for high-field mri: how much detail, effort, and accuracy is needed?* Magnetic resonance in medicine **69**, 1157 (2013).

[8] H. Homann, P. Börnert, H. Eggers, K. Nehrke, O. Dössel, and I. Graesslin, *Toward individualized sar models and in vivo validation,* Magnetic resonance in medicine **66**, 1767 (2011).

[9] A. Torrado-Carvajal, Y. Eryaman, E. A. Turk, J. L. Herraiz, J. A. Hernandez-Tamames, E. Adalsteinsson, L. L. Wald, and N. Malpica, *Computer-vision techniques for water-fat separation in ultra high-field mri local specific absorption rate estimation,* IEEE Transactions on Biomedical Engineering **66**, 768 (2018).

[10] L. Golestanirad, A. A. Rahsepar, J. E. Kirsch, K. Suwa, J. C. Collins, L. M. Angelone, B. Keil, R. S. Passman, G. Bonmassar, P. Serano, *et al.*, *Changes in the specific absorption rate (sar) of radiofrequency energy in patients with retained cardiac leads during mri at 1.5 t and 3t,* Magnetic resonance in medicine **81**, 653 (2019).

[11] M. Deng, R. Yu, L. Wang, F. Shi, P.-T. Yap, D. Shen, and A. D. N. Initiative, *Learning-based 3t brain mri segmentation with guidance from 7t mri labeling,* Medical physics **43**, 6588 (2016).

[12] P. A. Yushkevich, Y. Gao, and G. Gerig, *Itk-snap: An interactive tool for semi-automatic segmentation of multi-modality biomedical images,* in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE, 2016) pp. 3342–3345.

[13] I. Laakso, S. Tanaka, S. Koyama, V. De Santis, and A. Hirata, *Inter-subject variability in electric fields of motor cortical tdcs,* Brain stimulation **8**, 906 (2015).

[14] R. Kikinis, S. D. Pieper, and K. G. Vosburgh, *3d slicer: a platform for subject-specific image analysis, visualization, and clinical support,* in *Intraoperative imaging and image-guided therapy* (Springer, 2014) pp. 277–289.

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

[16] U.-S. Choi, H. Kawaguchi, Y. Matsuoka, T. Kober, and I. Kida, *Brain tissue segmentation based on mp2rage multi-contrast images in 7 t mri,* PloS one **14**, e0210803 (2019).

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

[18] C. M. Collins, B. Yang, Q. X. Yang, and M. B. Smith, *Numerical calculations of the static magnetic field in three-dimensional multi-tissue models of the human head,* Magnetic resonance imaging **20**, 413 (2002).

[19] W. Liu, H. Wang, P. Zhang, C. Li, J. Sun, Z. Chen, S. Xing, P. Liang, and T. Wu, *Statistical evaluation of radiofrequency exposure during magnetic resonant imaging: application of whole-body individual human model and body motion in the coil,* International journal of environmental research and public health **16**, 1069 (2019).

[20] E. A. Rashed, J. Gomez-Tames, and A. Hirata, *Deep learning-based development of personalized human head model with non-uniform conductivity for brain stimulation,* IEEE Transactions on Medical Imaging (2020).

[21] E. A. Rashed, J. Gomez-Tames, and A. Hirata, *Development of accurate human head models for personalized electromagnetic dosimetry using deep learning,* NeuroImage **202**, 116132 (2019).

[22] O. Ronneberger, P. Fischer, and T. Brox, *U-net: Convolutional networks for biomedical image segmentation,* in *International Conference on Medical image computing and computer-assisted intervention* (Springer, 2015) pp. 234–241.

[23] K. Yee, *Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media,* IEEE Transactions on antennas and propagation **14**, 302 (1966).

[24] X. G. Xu and K. F. Eckerman, *Handbook of anatomical models for radiation dosimetry* (CRC press, 2009).

[25] X. Xu, T. Chao, and A. Bozkurt, *Vip-man: an image-based whole-body adult male model constructed from color photographs of the visible human project for multi-particle monte carlo calculations,* Health Physics **78**, 476 (2000).

[26] J. Martin, C. Ashtiani, and M. Smith, *A finite element calculation of radiofrequency magnetic fields in simulated human organs and tissues,* in *Sixteenth Annual Northeast Conference on Bioengineering* (IEEE, 1990) p. 131.

[27] M. J. Ackerman, *The visible human project/sup tm: a resource for anatomical visualization,* in *Information Technology Applications in Biomedicine. ITAB'97. Proceedings of the IEEE Engineering in Medicine and Biology Society Region 8 International Conference* (IEEE, 1997) pp. 29–31.

[28] W. Liu, C. Collins, and M. Smith, *Calculations ofb 1 distribution, specific energy absorption rate, and intrinsic signal-to-noise ratio for a body-size birdcage coil loaded with different human subjects at 64 and 128 mhz,* Applied magnetic resonance **29**, 5 (2005).

[29] P. Dimbylow, *Fdtd calculations of the whole-body averaged sar in an anatomically realistic voxel model of the human body from 1 mhz to 1 ghz,* Physics in Medicine & Biology **42**, 479 (1997).

[30] P. Dimbylow, *Development of the female voxel phantom, naomi, and its application to calculations of induced current densities and electric fields from applied low frequency magnetic and electric fields,* Physics in Medicine & Biology **50**, 1047 (2005).

[31] Y. Shao, S. Shang, and S. Wang, *On the safety margin of using simplified human head models for local sar simulations of b1-shimming at 7 tesla,* Magnetic Resonance Imaging **33**, 779 (2015).

[32] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, and Y. Wang, *Artificial intelligence in healthcare: past, present and future,* Stroke and vascular neurology **2**, 230 (2017).

[33] A. Patel, *Chapter-7 under-fitting, over-fitting and its solution,* (2018).

[34] S. Pathan and P. Thakre, *Fuzzy clustering technique and pca based unsupervised change detection method in multitemporal sar images,* International Journal of Engineering and Management Research (IJEMR) **7**, 345 (2017).

[35] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, *N4itk: improved n3 bias correction,* IEEE transactions on medical imaging **29**, 1310 (2010).

[36] W. M. Brink, V. Gulani, and A. G. Webb, *Clinical applications of dual-channel transmit mri: A review,* Journal of Magnetic Resonance Imaging **42**, 855 (2015).

[37] F. Heckel, M. Schwier, and H.-O. Peitgen, *Object-oriented application development with mevislab and python.* (2009) pp. 1338–1351.

[38] S. M. Smith, *Fast robust automated brain extraction,* Human brain mapping **17**, 143 (2002).

[39] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, *User-guided 3d active contour segmentation of anatomical structures: significantly improved efficiency and reliability,* Neuroimage **31**, 1116 (2006).

[40] Y. Geng, Y. Ren, R. Hou, S. Han, G. D. Rubin, and J. Y. Lo, *2.5 d cnn model for detecting lung disease using weak supervision,* in *Medical Imaging 2019: Computer-Aided Diagnosis*, Vol. 10950 (International Society for Optics and Photonics, 2019) p. 109503O.