



UNIVERSITY OF AMSTERDAM

MSC ARTIFICIAL INTELLIGENCE
MASTER'S THESIS

Towards Open-Source MRI Data: Defacing and Sharpness Estimation

by

EMMA HOKKEN

10572090

June 30, 2021

Number of Credits: 48
September 2020 - June 2021

Supervisors:

C. ZHANG,
Dr. M.W.A. CAAN, and
Dr. X. ZHEN

Examiner:

Dr P.S.M. METTES



Abstract

Magnetic Resonance Imaging (MRI) is a widely used tool in both research and clinical settings, with image quality increasing every year. However, the acquisition time of a high-quality MRI can be very long. Accelerated MRI is a method in which this acquisition time is reduced by undersampling the available data points. Undersampling causes aliasing, i.e. overlap of regions, and adds noise to the image. A variety of methods are in place to reconstruct the image, with recent years showing a spike in reconstruction methods that utilise Deep Learning. One such method is the Recurrent Inference Machine (RIM).

This work comes in two parts. The first part aims to stimulate the research into Deep Learning-based accelerated MRI reconstruction by releasing the Amsterdam Ultra-high field adult lifespan database (AHEAD), a 7T dataset. To comply with privacy regulations such as GDPR, these images are to be defaced by the dilation of the brain mask. To simulate a real scan as much as possible, noise was added to the background. While the defacing method was able to accurately and precisely deface scans, the additional noise proved to be overestimated when compared to an original scan.

The second part focuses on creating a content-wise evaluation method for reconstructed MRI images. Currently, Deep Learning reconstruction methods are evaluated using quantitative measures. None of these evaluations take the image content quality into account. This work proposes a method that estimates the edge sharpness across tissue boundaries within an MR image using a two-fold k-means clustering approach. This method showed a decrease in sharpness for an increase in acceleration factor, showing that it is a valid method to evaluate image content quality and can provide valuable insights in Deep Learning-based accelerated MRI reconstruction.

Acknowledgements

I would first and foremost like to thank Chaoping Zhang, Matthan Caan, and Xiantong Zhen for their excellent supervision during this project. They helped me sort through ideas, come up with solutions, and gave excellent feedback on both coding and writing. Additionally, I would like to thank Pascal Mettes for agreeing to be the examiner for this master's thesis. I would also like to thank Pilou Bazin for providing initial data, his help in creating additional materials, and his suggestions and advice regarding the sharpness estimation. My thanks also go out to my colleagues in the Recon group at the Amsterdam UMC, Biomedical Engineering and Physics department, whose weekly discussions and journal readings broadened my knowledge and allowed for new and interesting insights into both this project and accelerated MRI reconstruction. In that regard, special thanks also go out to Kevin, for helping me set up and show me around the world of both MRI and the Amsterdam UMC. Thank you to my dad for his advice and support during this thesis, and I wish my mom could see the work I have delivered. Further, I would like to thank Martin for his incredibly useful comments on my writing, Karla for giving advice on making this thesis accessible for the colour blind, and Jordan, Julia, and Lin for their moral support during this project. You are the best rubber duckies out there. And finally, a silly thank you to Minoes, for keeping me company during my days at home and for reminding me when it's time to take a break by demanding scritches.

Contents

Abstract	i
Acknowledgements	i
List of Figures	iv
List of Tables	v
List of Abbreviations	vi
1 Introduction	1
1.1 Research Questions	2
1.2 Outline	2
2 Background	3
2.1 Magnetic Resonance Imaging (Westbrook et al., 2011)	3
2.1.1 Atoms and magnetism	3
2.1.2 Relaxation	3
2.1.3 Signal Acquisition	4
2.2 Image Reconstruction	5
2.2.1 K-space	5
2.2.2 The Fourier Transform	6
3 Related work	8
3.1 Standard reconstruction methods	8
3.2 Deep Learning reconstruction methods	9
3.3 Sharpness in Medical Imaging	10
4 Methods	11
4.1 AHEAD Dataset	11
4.2 Defacing Method	13
4.2.1 Defacing using Freesurfer	13
4.2.2 Defacing using dilation	13
4.2.3 Adding Gaussian background noise	14
4.2.4 Publishing AHEAD	14
4.3 Quantitative MRI Reconstruction using the RIM	14
4.3.1 RIM update equations	16
4.3.2 Update function	17
4.3.3 Loss function	17
4.4 Sharpness estimation along tissue boundaries	18
4.4.1 Distance function	18
4.4.2 Clustering	19
4.4.3 Coregistration	19
4.4.4 Subjects and ROIs	21
5 Results	22
5.1 Defacing	22
5.1.1 Defacing using Freesurfer	22
5.1.2 Defacing using dilation	22
5.1.3 Adding Gaussian background noise	25

5.2 Sharpness estimation along tissue boundaries	25
5.2.1 Trend analysis	28
6 Discussion	31
6.1 Defacing	31
6.1.1 Defacing method	31
6.1.2 Adding Gaussian background noise	31
6.1.3 Dataset publication	32
6.2 Sharpness estimation along tissue boundaries	33
6.2.1 Trend analysis	33
7 Conclusion	34
Bibliography	35
A Methods	38
B Results	39

List of Figures

2.1	Illustration of the generation of an MR signal.	4
2.2	Illustration of the relationship between points in k-space and points in image space. Each point in the image space maps back to a single point in k-space. Image inspired by mriquestions.com .	6
4.1	Representation of the multi-echo gradient echo sequence used to acquire the MRI data.	11
4.2	Four examples of scans that can be found in the AHEAD dataset that will be used in this work.	12
4.3	Image of example k-space data along phase-encoding dimensions depicting the ring used to calculate the standard deviation (STD) used to generate background noise.	15
4.4	2D single-slice depiction of the location of the cube that is used to compare the noise level in the original image and the defaced image.	15
4.5	An illustration of the network architecture of the RIM. Figure created by Zhang et al. (2021) .	17
4.6	Illustration of the clustering approach.	20
4.7	Outline of the used ROIs in the sharpness estimation.	21
5.1	Results of defacing using the Freesurfer defacing software for a single subject.	23
5.2	Freesurfer defacing applied to three different subjects.	23
5.3	Defacing using dilation of the brain mask for different numbers of iterations.	24
5.4	Defacing by dilation applied to three different subjects.	24
5.5	Defacing results using the dilation algorithm shown for different planes in one subject.	25
5.6	The percentage difference in noise from the original image compared to the newly generated noise of the defaced image.	26
5.6	The percentage difference in noise from the original image compared to the newly generated noise of the defaced image.	27
5.7	Comparison of the original image and the defaced scan.	27
5.8	The FWHM of one subject for different acceleration factors for both ground truth and accelerated image.	28
5.9	Profiles of the sharpness estimation showing whether a boundary is valid and contrasting enough to be included in the sharpness estimation.	29
5.10	Mean FWHM for all subjects per region for all acceleration factors.	29
5.11	Depiction of the distribution of sharpness in FWHM for all regions per acceleration factor.	30
A.1	Example of a coregistered levelmap used in the sharpness estimation.	38
B.1	Profiles from all acceleration factors of the sharpness estimation for the globus pallidus	40
B.2	Profiles from all acceleration factors of the sharpness estimation for the thalamus	41
B.3	Histograms for initial data inspection showing the distribution of FWHM data for two different acceleration factors.	42

List of Tables

4.1 Parameters of the scans included in the AHEAD dataset. *resulting from a built-in feature of the Philips system (Alkemade et al., 2020).	12
A.1 Distribution of ages of the participants from the AHEAD dataset. Numbers taken from Alkemade et al. (2020)	38

List of Abbreviations

AHEAD Amsterdam Ultra-high field adult lifespan database.

CNR contrast-to-noise ratio.

CS Compressed Sensing.

DFT Discrete Fourier Transform.

FFT Fast Fourier Transform.

FT Fourier Transform.

FWHM full width at half maximum.

GRAPPA Generalized Autocalibrating Partially Parallel Acquisitions.

IFFT Inverse Fast Fourier Transform.

IFT Inverse Fourier Transform.

MRI Magnetic Resonance Imaging.

NMV Net Magnetisation Vector.

PI Parallel Imaging.

PSF point spread function.

RF radio-frequency.

RIM Recurrent Inference Machine.

ROI region of interest.

SNR signal-to-noise ratio.

SSIM structural similarity index measure.

TE echo time.

TR repetition time.

Chapter 1

Introduction

Magnetic Resonance Imaging (MRI) is an important tool used in a variety of research and clinical settings ([Lønning et al., 2019](#)). MRI in its current form was developed in 1976 ([Rinck, 2019](#)). The predecessor of MRI as we know it today is known as Nuclear Magnetic Resonance (NMR). NMR describes the phenomenon that, when a small oscillating magnetic field is applied to atomic nuclei which have a strong magnetic field of their own, these nuclei will give off an electronic signal ([Hoult and Bhakar, 1997](#)). The very first NMR signals were taken of an anaesthetised rat in 1968 by [Jackson and Langham \(1968\)](#). Nine years later, the first MRI scan of the body of a human was taken by [Damadian et al. \(1976\)](#). Today, MRI is an important diagnostic tool and is widely used. As of 2018, there were approximately 50,000 MRI machines worldwide, with this number growing every year. In recent years, the sale of high field equipment, with magnet strengths of 1.5 Tesla or higher, have risen rapidly ([Rinck, 2019](#)). With the introduction of 7 Tesla scanners and its corresponding higher signal-to-noise ratio (SNR), diagnostic methods are improving continuously. However, the increased SNR allows for more complex acquisition and therefore also longer acquisition times ([Caan et al., 2019](#)).

Many studies have attempted to decrease these long acquisition times. This led to the development of Parallel Imaging (PI) ([Pruessmann et al., 1999](#)), as well as Compressed Sensing (CS) ([Lustig et al., 2007](#)). PI allows for reliable reconstructions with uniform undersampling of the measurements, while CS randomly samples data and reconstructs them into high-quality images by optimising over the L1 norm of coefficients in a sparsifying transform. In recent years, a new reconstruction trend has emerged: reconstruction using Deep Learning techniques. Deep Learning reconstruction methods can be split into two groups: iterative and non-iterative methods. Non-iterative methods take a holistic approach to reconstruction. They include networks such as the widely used U-Net by ([Ronneberger et al., 2015](#)). Iterative methods are more complex and work by updating the parameters of the network each iteration. Examples of iterative methods are the VarNet [Hammerlik et al. \(2018\)](#), ADMM-Net [Yang et al. \(2017\)](#), and the Recurrent Inference Machine ([Lønning et al., 2019; Zhang et al., 2021](#)).

While these Deep Learning methods have been shown to produce reliable results, they are not yet ready to be used in clinical settings. Usually, they are being evaluated and compared using quantitative measures such as the mean square error (MSE), normalised root mean square error (NRMSE), or structural similarity index measure (SSIM). These metrics are easy to compute and give great insight into the image quality but lack insight in terms of image content quality. To address this issue, an alternative method is to ask for reviews of clinicians. However, this is not only very time consuming, but opinions are subjective and may vary. Thus, a qualitative evaluation, which is able to access content-wise image quality of the reconstructions, would be beneficial. This work will focus on re-purposing a method of sharpness computation across tissue boundaries that can provide better insight into the content-wise image quality than metrics such as MSE or SSIM. The method works by using a two-fold k-means clustering method that is able to determine sharpness at the boundary of specific brain regions ([Bazin et al., 2020](#)).

Even though MRI scanners are widely available these days, they are still very costly to buy, with prices ranging from \$250,000 for low field machines to well above \$3,00,000 for high field machines ([Rinck, 2019](#)). Furthermore, the cost of a single MRI scan is also considerably high, with costs ranging from \$400 to \$5,000, depending on the body part and scan time. Besides that, an MRI scanner is space consuming. Because of this, not every hospital or research group has access to an MRI scanner. Additionally, high field scanners not only come with a monetary cost but also at the cost of patient comfort due to long scan times. In that light, this research will also partly focus on making a large, high-quality (7 Tesla) MRI dataset available to the public. The availability of this dataset will allow research groups without a scanner to access high-quality MRI data, but will also save researchers with access to an MRI scanner time by not having to scan. The main purpose of releasing this dataset is to help researchers improve accelerated MRI reconstruction, so that scan times can be reduced in the future.

Publishing the dataset comes with challenges of its own. We plan to release raw MRI data, which is often of

a large size due to the data being complex-valued and containing measurements for 32-channel radiofrequency receiver coils. Furthermore, there are restrictions in how data can be shared due to the privacy rights of the scanned subjects. To release this dataset and to comply with privacy rules such as GDPR ([Regulation, 2016](#)), these images must not contain any personal data that might lead back to the identity of the scanned person. Currently, the dataset that we aim to publish includes the faces of the participants. Before release, these faces should be removed. Part of this work will therefore focus on the anonymisation of this raw MRI data.

1.1 Research Questions

As mentioned above, this work consists of two parts. The first part focuses on the anonymisation of raw MRI data and the process of making it available to the public. While datasets with MR images are readily available, raw MRI data is commonly discarded due to storage capacity challenges and thus needs to be reacquired for this purpose. File size for raw MRI data can be up to 5 GB due to the data consisting of measurements from 32 coil elements and being complex-valued which makes it difficult to store. Another reason for raw MRI data not being widely available is that raw data contains information that MRI vendors (such as Philips or Siemens) would rather not have publicly available, as per non-disclosure agreements (NDAs). The release of this anonymous dataset will create opportunities for more people to get involved with research into accelerated MRI reconstruction, and possibly make cheaper, better, and more efficient MR imaging possible. Furthermore, it will allow Amsterdam UMC to host their own reconstruction challenges similar to [fastMRI](#) ([Zbontar et al., 2018](#)) in which research groups compete to create better MRI reconstruction using AI. Lastly, the International Society for Magnetic Resonance in Medicine (ISMRM) only allows such challenges that contain data from multiple vendors. Releasing this dataset will allow not only Amsterdam UMC but also others to host their customised challenges.

The second part of this work will address the qualitative evaluation of reconstructed MR images. When trying to increase the performance of accelerated MRI reconstruction models, it is wise to have multiple ways to evaluate results. As of now, there is not a structural, quantitative method of assessing content-wise reconstruction quality. There is a wide variety of statistical methods available (such as SSIM), but none of these methods focus on how the reconstructed image can be visually perceived by humans. This work will partly focus on creating a metric that will unveil insights into how well an (AI) method has performed reconstruction, content-wise, on an accelerated MRI. The more insight is available into the workings of certain reconstruction methods, the better one can work towards improving these methods. This will not only help individual research groups but can also be used in challenges such as [fastMRI](#) ([Zbontar et al., 2018](#)), which involve a range of research groups.

In summary, the objective of this work is to propose a process for publishing MRI data, making it available to a wider audience and for use in challenges. The two parts converge to the following research questions:

1. How can MRI data be defaced without compromising the raw data?
2. Is it possible to create a structural, quantitative method of assessing content-wise reconstruction quality by estimating the sharpness of tissue boundaries?

1.2 Outline

This work attempts to answer the above research questions. To do so, this work has been divided up into the following chapters.

- Chapter 2: Background. This chapter will present all background information needed to fully understand and make use of this thesis. It will focus on the physics of MRI and the tools needed to create an MR image.
- Chapter 3: Related Work. This chapter will explore prominent related work to the topics explored in this thesis; focusing on MRI reconstruction and sharpness evaluation in medical imaging.
- Chapter 4: Methods. This chapter will explain the methods used to answer the two research questions.
- Chapter 5: Results. This chapter will showcase the results from the created methods and experiments.
- Chapter 6: Discussion. This chapter will discuss the results, attempt to explain them, explore the limitations of this research, suggest future work, and work towards drawing conclusions.
- Chapter 7: Conclusion. This chapter will answer the research questions and conclude this work.

Chapter 2

Background

Over the course of this work, a certain amount of knowledge about the physics of Magnetic Resonance Imaging is assumed. In order to make this thesis available to a wider public, this assumed knowledge will be explored below.

2.1 Magnetic Resonance Imaging ([Westbrook et al., 2011](#))

2.1.1 Atoms and magnetism

As with most things, at the start of Magnetic Resonance Imaging (MRI) lie atoms. An atom consists of a central nucleus and electrons moving around this nucleus. While the nucleus is small, it contains most of the mass of an atom. This mass mainly comes from nucleons, which are divided into protons and neutrons. Protons have a positive electrical charge while neutrons have no charge. Electrons, on the other hand, lie outside of the nucleus, have a negative charge, and have only a fraction of the mass of a nucleus.

The parts of an atom are always in motion which can be classified into three different types. Electrons spin on their own axis, electrons spin around the nucleus, and the nucleus spins on its own axis. The principle of MRI is based on this last type of movement. The nucleus spins due to the individual movement of protons and neutrons. A pair of subatomic particles, i.e. a paired up proton and neutron, spin in opposite directions by default. One way to characterise atoms is to do so by their mass number. This number is the sum of protons and neutrons present in the nucleus. With atoms that have an even mass number, half of these pairs spin up in one direction, and the other half spin down in the other. However, in nuclei that have an odd mass number, this spinning is not balanced, causing an angular momentum in these nuclei. This attribute can be utilised in MRI scanning, as it causes the nuclei to align themselves with an applied magnetic field. These nuclei are therefore called MR active nuclei.

MRI also finds one of its fundamentals in Faraday's law of electromagnetic induction. This law refers to motion, magnetism, and charge and it states that if two of them occur together, they automatically induce the third. Nuclei that have a net charge and are spinning will therefore acquire a magnetic moment. Hydrogen nuclei are one example of this, as they consist of a single spinning, positively charged proton. Based on the law of electromagnetic induction, these nuclei obtain a magnetic moment. These hydrogen nuclei can thus act as a small magnet, with both a south and a north pole. When there is no external magnetic field, these nuclei are oriented randomly. However, when a strong external magnetic field is applied, they will align themselves with it. In MRI, this static external magnetic field is called B_0 . The majority of nuclei align themselves parallel (spin-up) with this field while a smaller number align themselves in an anti-parallel fashion (spin-down). However, as these individual magnetic moments are small, they cannot be used in clinical MRI. Instead, the Net Magnetisation Vector (NMV) is used. This vector reflects the ratio between spin-up and spin-down nuclei.

The influence of B_0 produces an additional spin in the hydrogen nuclei, called precession. This spin causes the nuclei to follow a circular path around the B_0 field. The value of this precession is determined by the Larmor equation:

$$\omega_0 = B_0 \times \lambda, \quad (2.1)$$

where λ represents the gyromagnetic ratio - the relationship between the magnetic moment of each MR active nucleus and the angular momentum.

2.1.2 Relaxation

During an MRI scan, a radio-frequency (RF) pulse is applied to the scanning target. When applying an RF pulse at the same rate as the Larmor frequency of a specific type of nucleus, these nuclei resonate with the B_0

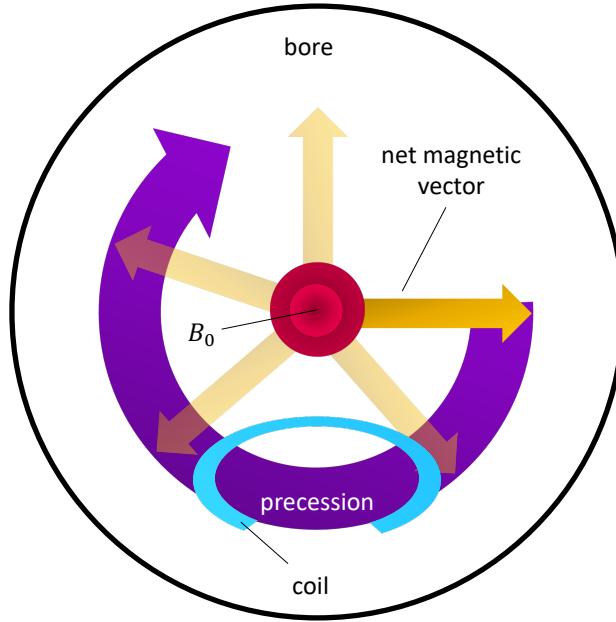


Figure 2.1: Illustration of the generation of an MR signal. B_0 is a static external magnetic field. When a radio-frequency pulse is applied, the net magnetic vector (NMV) gains energy and spins out of alignment with the B_0 field. When the RF pulse is turned off the NMV loses its energy and realigns with B_0 by following the precessional path to do so. A voltage can be measured in the coil when the NMV passes through the coil. Inspired by Figure 1.13 from [Westbrook et al. \(2011\)](#).

field. This process is called excitation. When resonance occurs, the nucleus gains energy from the pulse and the NMV moves out of alignment and away from B_0 as a consequence. The angle at which the NMV has moved out of alignment with the B_0 field is called the flip angle. Once the RF pulse is switched off, the NMV will return to its base state. In the process of realigning with the B_0 field, the hydrogen nuclei release the energy absorbed from the RF pulse. This, in turn, is called relaxation, during which the amount of magnetisation in the longitudinal plane (B_0) increases, i.e. the magnetisation recovers. At the same time, the amount of magnetisation in the transverse plane decreases, i.e. decays. This recovery of longitudinal magnetisation is called T_1 recovery, while the decay of transverse magnetisation is called T_2 decay.

When nuclei release their energy to the surrounding environment during the T_1 recovery, the rate of this energy emission is exponential and is called the T_1 recovery rate. Because of this exponential aspect, it is not necessary to wait until all energy has been released. Rather, it is sufficient to wait until 63% of the longitudinal magnetisation has been recovered in the tissue. While T_1 relaxation is also known as spin-lattice relaxation, T_2 decay is known as spin-spin relaxation. T_2 decay is caused by the fact that magnetic fields of neighbouring nuclei interact with each other - when T_1 relaxation occurs, T_2 relaxation also automatically occurs. In essence, T_2 decay is the loss of coherent magnetisation in the transverse plane. T_2 relaxation time refers to the time it takes 63% of the transverse magnetisation to be lost.

2.1.3 Signal Acquisition

The MR signal itself is produced by resonance and is based on the law of electromagnetic induction. The magnet of an MRI is located in the bore - the hollow centre of the machine where the scanner subject lies. The MR signal is produced when the magnetisation produced by the (hydrogen) nuclei cuts across a receiver coil. An overview of how the MR signal is created can be found in Figure 2.1.

An MRI scan consists of a sequence of RF pulses. An RF pulse sequence can be characterised by multiple time periods, with the two main periods being repetition time (TR) and echo time (TE). TR refers to the time between two consecutive RF pulses and therefore determines how much relaxation is allowed to occur. TE relates to the time it takes the signal to reach its peak after the excitation RF pulse and controls how much decay of transverse magnetisation is able to occur.

To create an image that shows different tissues in the body, it is useful to introduce the use of contrast. There are two different types of contrast: intrinsic and extrinsic contrast. Intrinsic contrast is decided by T_1 recovery time and T_2 decay time while extrinsic contrast parameters are TR, TE, flip angle, inverse time (T_1), and others. Contrast is created by the difference in signal strength from different tissues. As an example, if some tissue has a large transverse component of coherent magnetisation at time TE, it will return a high signal

and will appear bright on the resulting image.

While different tissues have different energy absorbing capabilities, they also lose their energy at different rates. This results in a difference in the T_1 and T_2 signal between tissues. Take the difference between fat and water, for example. Water cannot easily absorb energy and therefore has less longitudinal magnetisation before the application of an RF pulse than fat. After an RF pulse, water has less transverse magnetisation, therefore less signal and appears darker on a T_1 weighted image. The opposite holds for T_2 weighted image. Because the molecules of water are spaced further apart, T_2 decay is less efficient and because the decay is slower, water appears brighter than fat on a T_2 weighted image.

A third intrinsic contrast method is called T_2^* , which is typically imaged with a sequence called gradient echo (GRE). T_2 weighted images are hard to reproduce. This is partially due to the fact that a magnet in an MRI always has some inhomogeneities, meaning that the magnetic field does not have precisely the same strength everywhere in the bore. This results in a distortion of the signal. The T_2^* contrast takes this into account, so it is a combination of T_2 decay and dephasing due to these inhomogeneities. This dephasing results in a loss of coherent transverse magnetisation, i.e. signal. This means that often, the signal is gone before most tissues have had time to reach their T_1 or T_2 relaxation times. In order to counteract the field inhomogeneity induced relaxation, the signal needs to be regenerated. This can be done by either introducing an additional pulse or a gradient that turns the hydrogen nuclei 180° . This gradient allows for the signal decay not induced by T_2 relaxation to cancel out so the tissues have more time to reach their relaxation times. Using an additional pulse is referred to as spin echo pulse sequences, the method using a gradient is called gradient echo pulse sequences. The latter method will be further explored.

The initial RF excitation pulse used in a gradient echo pulse sequence varies and does not necessarily have to be 90° . After the excitation pulse, T_2^* dephasing immediately occurs due to an applied gradient. After this dephasing, another gradient with the opposite polarity is applied to rephase the signal. A gradient dephases signal by altering the magnetic field strength as experienced by the coherent magnetisation. As a consequence, a proportion of the magnetic moments speed up while others slow down depending on the position on the gradient axis of the moments. Rephasing using a gradient functions similarly: the magnetic field strength is altered, which ensures that the slower nuclei speed up due to their position on the gradient axis. In essence, gradients are an alteration to the main magnetic field.

Using a gradient echo pulse sequence has the advantage of both a shorter TE and TR. TE can be shorter because gradients rephase faster than RF pulses with a 180° flip angle. TR can be reduced partially because TE is reduced, but mainly because of the variable flip angle used in gradient echo pulse sequences. A lower flip angle ensures faster recovery of longitudinal magnetisation and thus shorter TR. A downside of using a gradient echo pulse sequence is that it is not possible to compensate for field inhomogeneities. Because of this, a gradient echo sequence often leads to T_2^* weighted imaging.

In this work, the main focus will be on scans that use a R_2^* contrast. The R , in this case, refers to the relaxation rate. T_1 , T_2 , and T_2^* are relaxation times. These can be converted into relaxation rates, R_1 , R_2 , and R_2^* , respectively. These relaxation rates are the inverse of the relaxation times, so $R_2 = 1/T_2$. This also means that, while T1 is usually measured in ms, R1 is measured in 1/ms.

2.2 Image Reconstruction

2.2.1 K-space

MR data is collected by the frequency at which the magnetic moments cross the receiver coil and their position or phase on their precessional path (Westbrook et al., 2011). This information is combined into one data point, which is stored in so-called k-space. K-space is a spatial frequency domain, meaning that it stores the origin or phase of the signal and the frequency of the signal. K-space stores information of the frequency on the x-axis and information about the phase on the y-axis. The axes in k-space are numbered differently than one might be used to. The origin of the axes lies at the centre of k-space, with the outer lines of the axes lying near the edge of k-space. The top half of k-space, meaning the first and second quadrant, will contain lines with positive numbers, indicating a positive polarity, and the bottom half of k-space, the third and fourth quadrant, will contain lines with negative indices, indicating a negative polarity. It is important to note that k-space does not manifest as the image itself: the image is represented in a transformed k-space called image space. Furthermore, there is no spatial correspondence from k-space to image space, as every data point in k-space contributes information to every pixel in the image (Mezrich, 1995). Figure 2.2 shows an arbitrary example of how points in k-space relate to signal frequency in image space. The transition from k-space in the frequency domain to the image itself in the image domain is done by applying the Fourier Transform, often implemented as the Fast Fourier Transform.

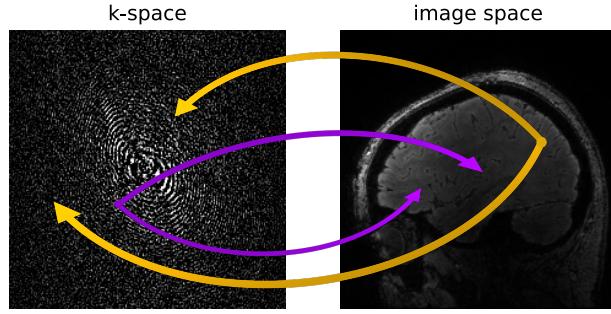


Figure 2.2: Illustration of the relationship between points in k-space and points in image space. Each point in the image space maps back to a single point in k-space. Image inspired by mriquestions.com.

2.2.2 The Fourier Transform

The Fourier Transform (FT) is a mathematical operation that decomposes a function into its separate frequencies. Bernstein et al. (2004) provides the analogy of a piano. When applying an FT to a chord played on a piano, the FT will show which keys were struck and how hard they were struck. In the case of MRI, the FT transforms data that is in the image domain to the frequency domain. The FT is defined as follows:

$$FT(g(x)) = G(k) = \int_{-\infty}^{+\infty} g(x) \exp^{-2\pi i k x} dx \quad (2.2)$$

Here, x and k are real variables known as Fourier conjugates (Bernstein et al., 2004). They represent a pair of Fourier domains, which in turn are usually time and frequency or distance and k-space.

To reconstruct MR images, the Inverse Fourier Transform (IFT) is used. It transforms data by decomposing the combined frequencies of the signal into singular peaks at a specific frequency (Bernstein et al., 2004), i.e. it transforms the data from the frequency domain to the image domain. The IFT is defined as follows:

$$IFT(G(k)) = g(x) = \int_{-\infty}^{+\infty} G(k) \exp^{+2\pi i k x} dk \quad (2.3)$$

Due to the way data is sampled in MRI, the FT is not fully necessary as it assumes a function of a continuous variable (Bernstein et al., 2004). Therefore, in MRI reconstruction, mainly the Discrete Fourier Transform (DFT) is used. Given a string of N complex data points, $d = d_0, d_1, d_2, \dots, d_{N-1}$ the J th element of the DFT is defined as

$$DFT(d)_J = D_J = \sum_{K=0}^{N-1} d_K \exp^{\frac{-2\pi i J K}{N}} \quad (2.4)$$

with the inverse DFT being

$$IDFT(D)_K = d_K = \frac{1}{N} \sum_{J'=0}^{N-1} D_{J'} \exp^{\frac{+2\pi i J' K}{N}} \quad (2.5)$$

Here, the $\frac{1}{N}$ is required for normalisation, to ensure that the values can be mapped back to each other. N refers to the number of data points.

However, the DFT has an operational time complexity of $O(N^2)$ (Bernstein et al., 2004). With MRI scanners required to perform a large number of DFTs, this complexity is not feasible. Because of this, when translating from k-space to image space, the Fast Fourier Transform (FFT) is applied instead. The FFT is an implementation of the DFT and yields the same result, but does so in $O(N \log N)$ which makes it more computationally efficient. There are several algorithms available to accomplish this, but at its core, the FFT splits the problem into multiple, smaller subproblems. If the number of points N is even, the signal can be split into two sub-signals. The DFT can be expressed as a linear combination of DFTs of these two sub-signals. This can be repeated recursively until a DFT is calculated for each point individually. Of course, the FFT also has its inverse counterpart named IFFT.

In practice, an FFT only is not sufficient to create a proper MR image. Most FFT implementations assume that the origin of k-space lies in the same position as the origin of image space. Because the origin of the axes in k-space lies in the centre and the origin of image space does not, a shift needs to be applied. Luckily, most toolsets provide a so-called FFTshift that shifts the origin from the centre to the second quadrant. IFFTshift will undo this shift. To reconstruct an MR data d into an image i , the following order of operations is applied:

$$i = \text{IFFTshift}(IFT(FFTshift(d))) \quad (2.6)$$

The above explanation has thus far assumed one-dimensional data. For MRI, this is unrealistic as most MR data is at least 2D. Luckily, there exist multi-dimensional versions of all variations of the FT. When reconstructing MR data, this FFT is applied in all spatial dimensions.

Chapter 3

Related work

MRI typically has a long acquisition time, often making it impractical or uncomfortable for patients or research. Thus, the need to decrease acquisition time arose. Ever since the 1970s, new methods have been developed to increase imaging speed, both with respect to hardware, e.g. improved magnetic field gradients, and software, e.g. new pulse sequences (Zbontar et al., 2018).

Initial attempts at decreasing MR scan time were mainly focused on increasing the number of receiver coils (Hutchinson and Raff, 1988; Kwiat et al., 1991). The multiple coils sampled signals from the same part of k-space, each with different coil sensitivities. These multiple sets of measurements allowed for interpolation over unsampled k-space data, thus decreasing acquisition time. However, according to Pruessmann et al. (1999), these attempts at reducing scan time were, at the time, not yet applied in clinical settings.

When undersampling k-space to decrease acquisition time, the data must be reconstructed to obtain a usable image. Reconstructing undersampled MR images is in essence an inverse problem. The forward model, how the true image is translated to the measured signal (e.g. by applying the FT), is known. What is not known, is how the measured signal can be transformed back into the image space to reflect the true image (Lønning et al., 2019). This is the case because k-space is undersampled, not allowing for the complete information of the full image. Reconstructing an MR image is not an easy task: the number of possible reconstructions is endless even when the area that has been scanned is known. Especially since one must also take into account potential image corruptions. Because of this, combinatorial ways of reconstruction are infeasible. Therefore, alternative reconstruction methods are needed.

3.1 Standard reconstruction methods

One method of reducing MR scan time is Parallel Imaging (PI). In short, PI allows for multiple sets of data points to be sampled simultaneously with the multiple receiver RF coils. If one knows how sensitive a certain receiver is to certain spatial positions, information about where an MR signal originated from can be inferred. Sensitivity is a property of the receiver itself and is therefore not dependent on the scanned sample (Pruessmann et al., 1999). The spatial sensitivity of different receiver coils within the scanner can be known beforehand from pre-scanning or computation from auto-calibration signals, and therefore can provide information that is necessary to interpolate undersampled signals. When reconstructing an image from undersampled k-space data, the image will contain overlap (wrap around) artefacts known as aliasing. By utilising knowledge about spatial sensitivity, an undersampled MR image can be dealiased.

The first practical application of parallel imaging to reduce scan time was made when SiMultaneous Acquisition of Spatial Harmonics (SMASH) was introduced (Sodickson and Manning, 1997). Scan time is reduced by making efficient use of the spatial information from a surface coil array. SMASH functions by acquiring signals from multiple surface coils simultaneously and using linear combinations of these signals to generate multiple datasets. These datasets are distinctly offset in k-space, allowing for the combining of these datasets into a complete k-space. The reconstruction is done before applying the Fourier transform. However, this method is restricted to certain combinations of coil arrangements, slice geometry, and acceleration factor (Pruessmann et al., 1999).

Pruessmann et al. (1999) introduced a method to overcome these restrictions called sensitivity encoding (SENSE). Prior to the MRI scan, the coil sensitivities are calculated. During the scan, only every n th line of k-space ($n = \text{acceleration factor}$) is acquired. Then the k-space data is translated to image space using the Fourier Transform. A matrix inversion is then applied to unfold these aliased points. SENSE is offered on a wide range of MRI machines and is, therefore, one of the most used reconstruction methods (Blaimer et al., 2004).

Another parallel imaging method is Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA)

(Griswold et al., 2002). This method is similar to SENSE, except that it corrects the aliasing before applying the Fourier Transform to generate images. With GRAPPA, the coil weighting factors are known from the auto-calibration signals acquired in k-space. These weighing factors include information on how much each coil distorts, smears, and displaces k-space data. With this information known, the missing k-space data is estimated. While GRAPPA might be less available than SENSE, it is especially useful in lung and abdominal MRI or when obtaining a coil sensitivity map is a difficult task (Blaimer et al., 2004).

In 2006, Compressed Sensing (CS) was introduced (Lustig et al., 2007). Different from the previously mentioned methods, CS uses randomly sampled k-space data to reconstruct the images. Reconstruction is done by optimisation over the L1 norm of coefficients of a sparsifying transform, which leads to state of the art undersampling factors of that time.

Although CS has many advantages over previous methods, it also has drawbacks. CS can require long reconstruction times, and its performance is dependent on the selection of the right hyperparameters. When the wrong hyperparameters are chosen, the image either becomes too smooth because of over-regularisation or still contains undersampling artefacts (Hammerink et al., 2018). These drawbacks motivate the emerging Deep Learning based methods which can largely solve these issues.

3.2 Deep Learning reconstruction methods

Although the above-mentioned methods yield good results for methods such as CS or PI, each new reconstruction is treated as a whole new optimisation problem, whereas most MR images are structurally similar to each other (Hammerink et al., 2018). Deep Learning-based methods, on the other hand, can make use of this similarity. There are two main Deep Learning methods for reconstructing MR images: iterative methods and non-iterative methods.

An iterative method was proposed by Hammerink et al. (2018). This method learns the key parameters of the inverse transformation. It uses a Variational Network that learns a complete reconstruction procedure for multi-channel MR images. This way, the problem is tackled as if it was a reaction-diffusion process as described in Chen et al. (2015). The MR reconstruction is treated as a variational model and is placed in a gradient descent learning scheme. Another method named ADMM-Net is based on the alternating direction method of multipliers (ADMM) approach for optimising CS-MRI. It uses dummy variables to solve the reconstruction problem Yang et al. (2017). This takes place through a series of partial updates. This model attempts to learn all parameters for MR images, including image transforms and shrinkage functions. In 2017, another iterative method by Schlemper et al. (2017) was introduced which uses a combination of cascading Convolution Neural Networks (CNNs) and additional layers to ensure data consistency. This data consistency layer corrects any ill-placed k-space data. The method, at the time, outperformed state of the art CS methods and proved that CNNs are able to learn spatio-temporal correlations efficiently by combining convolution and data sharing approaches.

As mentioned above, not all methods use an iterative approach. One such example is the U-net architecture as presented by Ronneberger et al. (2015). U-nets were initially designed for image segmentation, but have since been altered to handle inverse problems in both CT and MRI (Jin et al., 2016; Hyun et al., 2018). U-nets consist of two parts: feature extraction and spatial context creation. In the first part, convolutional layers, combined with max-pooling layers, are used to extract features from local patches in the input image. At the same time, the number of feature maps is steadily increased. This allows for the extraction of a large number of features, but also induces the loss of global context. Such a global context is precisely what is needed for reconstruction. The second part of a U-net uses unpooling layers to redress this loss. The unpooling layers are used to upsample the extracted feature maps to the same size as the original input image. At each unpooling layer, previously extracted feature maps at the current resolution level are concatenated with the output of the unpooling layer. This enables the network to combine the feature map with the spatial context. Although this non-iterative solution produces good reconstruction results, the network has the chance to easily overfit on the training data (Lønning et al., 2019).

Another non-iterative method is automated transform by manifold approximation (AUTOMAP) (Zhu et al., 2018). This reconstruction method learns a mapping from the sparsely sampled k-space input directly to the fully sampled image space output. Because of this mapping, the network learns an implicit low-dimensional manifold of the data for both the image domain and the k-space domain. This manifold is then conditioned over the reconstruction function. This joint manifold captures a highly expressive representation of the image and is robust to noise. However, this does require an IFT to be applied in order to map each point in k-space to all points in image space which in turn requires fully connected layers (Lønning et al., 2019).

A final reconstruction method is the Recurrent Inference Machine (RIM) by Lønning et al. (2019). The RIM is an iterative approach, which, due to its recurrent architecture, can share parameters across iterations. At the core of the RIM stands that it learns an inference algorithm given data and a task. This removes the need for knowledge about the domain, allowing a network architecture to learn a reconstruction method without prior

knowledge (Putzky and Welling, 2017). Because the RIM operates in the solution space, it makes for an efficient approach. Furthermore, as the RIM is based on a Recurrent Neural Network (RNN), it is Turing Complete and can therefore be trained to implement any inference algorithm (Siegelmann and Sontag, 1991; Putzky and Welling, 2017). The model used to reconstruct the MRI data in this work was introduced by Lønning et al. (2019) and updated by Zhang et al. (2021) and will be further discussed in Section 4.3.

3.3 Sharpness in Medical Imaging

The quality of images is often indicative of the mechanisms producing them. For biomedical images, there are a variety of methods available to determine image quality, e.g. signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR), artefacts, and edge sharpness. While the first methods are widely used, edge sharpness is not yet a prominent measure in MRI (Ahmad et al., 2015). Because MRI reconstruction methods offer a trade-off between image sharpness and SNR, image sharpness is a good method to quantify the reconstructed image.

In biomedical imaging, edges and their sharpness play an important role, as they contain physiologically important information. Various attempts to create metrics for edge sharpness have been proposed. One such method relies on computing local gradient values of the image intensity, i.e. rise and fall of image intensity (Rieger and van Veen, 2008; Biasioli et al., 2011). However, MRI images can be subject to low SNR, and this gradient method is quite sensitive to noise. Furthermore, this method might not be suitable for the brain, as it requires objects that have a constant intensity, something which the complex-structured brain does not always have. Another method makes use of local edge kurtosis (Caviedes and Gurbuz, 2002) but is also quite sensitive to noise.

Besides the aforementioned issues, finding a sharpness measure for MR images, and specifically the brain, brings forth two more challenges. In the brain, myelination and iron deposition patterns can vary for different brain regions (Lebel et al., 2012). As a consequence, a metric that uses quantitative parameters will not result in a trustworthy sharpness measure. Furthermore, the tissue signal may not be strong enough, so only part of the tissue boundary will show enough contrast to calculate the sharpness (Bazin et al., 2020). Thus, a sharpness measure must be resistant to both of these challenges.

Ahmad et al. (2015) proposed a sharpness measure they coined Edge Sharpness Assessment by Parametric modelling (ESAP). ESAP allows for the estimation of image sharpness for specific regions of interest (ROIs) and consists of four steps: edge selection using a canny edge detector (CED), intensity profile extraction, least-squares curve fitting, and hypothesis testing. While this method gives better insights than SNR, CNR, local gradients or edge kurtosis, it still requires the user to select the ROI and set the CED parameters.

Bazin et al. (2020) defined a sharpness measure in terms of the transition boundary between an ROI and the surrounding tissue. This can be done for both cortical and subcortical structures. It uses a sigmoidal function to model the intensity transition and fit the sharpness estimation of the ROIs' edge to signal values. This allows for an accurate estimation of the edge width and its corresponding uncertainty. A similar sharpness measure was used by Schoormans et al. (2020). They assessed image sharpness by fitting a sigmoidal function perpendicular to the surrounding air/tissue boundary. However, because this was done using the air/tissue boundary, this method might not work well for the complexity that MRI scans of the brain bring forth.

Since the sharpness estimate of Bazin et al. (2020) is defined in terms of physical distance, this method is invariant to the scaling of input signal values. Because the brain is a complex organ and variability in the tissue signal is likely, the sharpness estimate is followed by a clustering approach to allow for sharpness definition within sub-regions of similar tissue. Further, the sharpness estimate is only performed when enough contrast with the surrounding tissue is available. If the measure does not have enough contrast, an estimate cannot be made for that specific ROI and will ignore it. Whether or not enough contrast is present is assessed by the model fit quality. This method by Bazin et al. (2020) will be used in this work, and thus a more detailed description of this method can be found in Chapter 4.

Chapter 4

Methods

4.1 AHEAD Dataset

The main dataset that was used in this work was the Amsterdam Ultra-high field adult lifespan database (AHEAD) ([Alkemade et al., 2020](#)). This dataset consisted of 105 7T whole-brain structural MRI scans. These scans were adapted specifically to image the human sub-cortex, but the set also included scans of the cortex. The dataset contained 105 scans from both female and male participants, ranging from 18 to 80 years old, which were divided into six age groups. Each age group contained at least six male and six female participants, with also six male and six female participants in each decade. Participants in the youngest age group were scanned specifically for the creation of this dataset, while other age groups were scans taken at other times, but with the same sequence, within the same research group. A precise makeup of the demographic can be found in Appendix A, Table A.1.

Scans were made at the Spinoza Centre for Neuroimaging in Amsterdam, the Netherlands. A Phillips Achieva 7T MRI scanner with a 32-channel head array coil was used and images were taken using the multi-echo magnetisation-prepared rapid gradient echo (MP2RAGEME) sequence ([Caan et al., 2019](#)). This sequence is an extension of the earlier created MP2RAGE sequence by [Marques et al. \(2010\)](#). The MP2RAGEME sequence consisted of two rapid gradient echo images that were acquired in the sagittal plane after an inversion pulse of 180° and inversion times of $TI_1 = 670ms$ and $TI_2 = 3675.4ms$. The second inversion was extended by adding a multi-echo readout at four echo times $TE_1 = 3ms$, $TE_{2,1-4} = [3, 11.5, 19, 28.5]ms$. Both flip angles were $FA_{1,2} = [4^\circ, 4^\circ]$. Scans were taken with an acceleration factor of 2 and were reconstructed using GRAPPA ([Griswold et al., 2002](#)). All parameters are listed in Table 4.1. The turbo factor of 150 resulted in 176 shots. The total acquisition time was 19.53 minutes. An illustration of the second echo time acquisition is depicted in Figure 4.1.

From these scans, R_1 -maps, R_2^* -maps, T_1 -maps, T_1 -weighted images, T_2^* -maps, and Quantitative Susceptibility Mapping (QSM) images were derived. Because contrasts were derived from a single scan, co-registration was not needed. Images were reoriented to be in line with radiological display convention. Participant motion was limited in the 105 images included in the dataset, so these images can be used for sharpness estimation in which it can be assumed that motion does not play a significant effect. Data was converted to NIFTI format by [Alkemade et al. \(2020\)](#). An example of the data and the different scans can be found in Figure 4.2. The dataset also contains complex coil sensitivities of the scans.

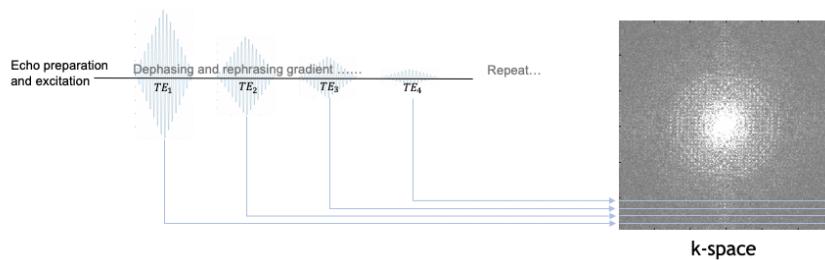


Figure 4.1: Representation of the multi-echo gradient echo sequence used to acquire the MRI data.

Parameter	Value
Inversion pulse	180°
TI_1	670 ms
TI_2	3675.4 ms
TE_1	3 ms
$TE_{2,1-4}$	3, 11.5, 19, 28.5 ms
$FA_{1,2}$	[4°, 4°]
$TR_{GRE1,2}$	[6.2 ms, 31 ms]
bandwidth	404.9 MHz
$TR_{MP2RAGEME}$	6778 ms
acceleration factor	2
FOV	205 × 205 × 164 mm
acquired voxel size	0.7 × 0.7 × 0.7 mm
acquisition matrix	292 × 290 × 234
reconstructed voxel size*	0.64 × 0.64 × 0.7 mm
turbo factor (TFE)	150

Table 4.1: Parameters of the scans included in the AHEAD dataset. *resulting from a built-in feature of the Philips system ([Alkemade et al., 2020](#)).

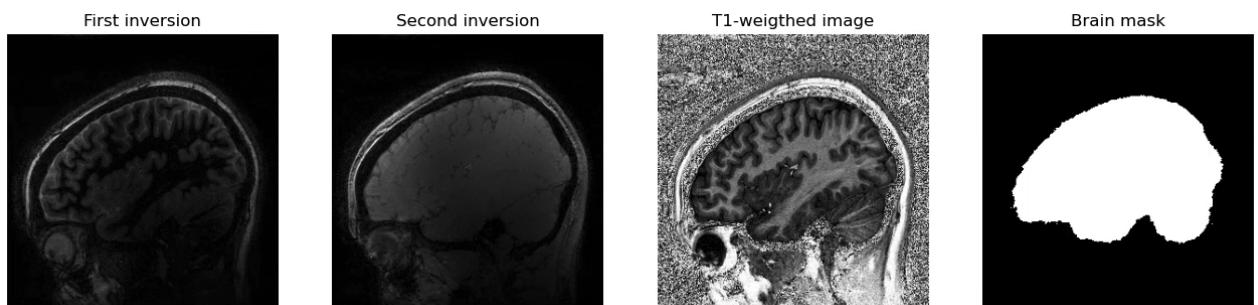


Figure 4.2: Four examples of scans that can be found in the AHEAD dataset that will be used in this work. All scans originate from one subject.

4.2 Defacing Method

In order to release the AHEAD dataset to the public, the MRI images need to remain anonymous. Because regions of the face are visible, this requires some pre-processing. To anonymise the data, two defacing methods were investigated: defacing software from Freesurfer (Bischoff-Grethe et al., 2007) and a dilation algorithm.

4.2.1 Defacing using Freesurfer

The first method was the defacing software from Freesurfer (Bischoff-Grethe et al., 2007). To remove facial structures, the algorithm used a model of non-brain structures. The facial features of ten subjects were manually labelled to create an atlas of face membership. Facial features comprised the entire front of the head. To be able to remove facial features from never-before-seen scans, brain and non-brain structures were used as input to an optimal linear transform. After this, a brain mask was constructed. This was done by forming a union with all voxels that had a nonzero prior probability of being any brain tissue. From this mask, a binary volume was created by morphologically dilating the mask n times ($n = 7$). Nonzero values of this volume indicated the presence of brain tissue within nx millimetres. Here, x represents the size of a voxel in milometers. The volumes were interpolated to ensure isotropic voxel dimensions. The $n = 7$ dilations ensured that nonzero probabilities were preserved. This way the accuracy of the linear transform could be guaranteed.

The deidentification involved finding all voxels that were outside of the mask and also having a nonzero probability of being a facial feature. Voxels within x or $2x$ of the brain mask were removed if the Mahalanobis distance to any brain tissue was low. To calculate this distance, mean and covariances were estimated using a manually labelled training set. This last part proved to be particularly useful when removing fatty regions around the orbital areas. Because the defacing software was created using T_1 -weighted images, the authors note that it should only be used on T_1 -weighted images.

Initial analysis of the defacing software showed that it did not always fully remove the face completely at times, likely due to the noisy background of the input image. Therefore, the Nighres package was used to determine what part of the image was the background and what part was the head (Huntenburg et al., 2018). The background was calculated using the scan of the second inversion, after which the resulting background mask was applied to the defacing input image. The Freesurfer defacing tool also occasionally removed parts of the brain. To rectify this, the mask of the brain was used to ensure removed brain regions were put back.

Lastly, the transition boundary between the removed face and the background was quite sharp. This is very much unlike the rest of the image and can be problematic when analysing the MRI data further. Because of this, a Gaussian filter was applied to the defaced image along the face boundary. A Gaussian filter is a way of smoothing an image by applying a convolution on the image using a Gaussian kernel (Szeliski, 2010). Formally, a Gaussian filter follows this equation:

$$\mathcal{G}(x, y; \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (4.1)$$

For the application of the filter, a mask of the removed face was created and the Gaussian filter applied to this mask. The smoothed mask was then multiplied with the defaced image to create a smoother boundary. The kernel size was set to five.

4.2.2 Defacing using dilation

The second method that was investigated was dilation of the brain mask. Because the brain mask is a binary image, the definition of binary dilation can be used, and reads as follows:

$$A \oplus B = \{a + b : a \in A, b \in B\} = \bigcup_{b \in B} A_b \quad (4.2)$$

The second equality holds if A_b is defined as $A_b = \{a + b : a \in A\}$ (Wilson and Ritter, 2000).

Initially, the idea was to use a relatively small kernel and perform multiple iterations of this to ensure more precise dilation. Using a smaller kernel for multiple iterations could more easily show how much dilation was required. If needed, a bigger kernel could be used to achieve the equivalent dilation with 1 iteration. It was theorised that using a bigger kernel would be more computationally efficient, as applying a matrix operation only once is more efficient than applying a smaller matrix operation multiple times. For example, mathematically, dilating with a 3×3 kernel for 2 iterations is the same as dilating with a 5×5 kernel for 1 iteration given that the 5×5 kernel is a dilated equivalent to the 3×3 kernel (Jain, 1989). However, testing with a $50 \times 50 \times 50$ significantly slowed down the dilation when using the Scipy package, going against the expectation. Because there is no mathematical difference in dilating with a bigger kernel for one iteration and dilating with a smaller kernel for more iterations, a smaller kernel was used.

In this specific instance, dilation on the binary brain mask was performed using a $5 \times 5 \times 5$ kernel and applied for 10 iterations. The kernel size and number of iterations were generated after some experiments which are further described in Section 5.1.2. Dilating the brain mask ensured that all brain regions were included in the defaced image. Therefore, it is not necessary to include a check into the algorithm for this.

Just like with the Freesurfer defacing method, the edge region of the defaced image needed to be smoothed out. For this, the same Gaussian smoothing was applied as described in Equation 4.1. The brain mask was dilated, smoothed, and then placed over the scan that needed to be defaced. Because the brain mask is a binary image, the background of the defaced image was removed. In order to create a defaced image that is as close to the original as possible, the background was filled with simulated noise. A few approaches to adding noise were considered and explained in the following section.

4.2.3 Adding Gaussian background noise

A first method to gauge noise levels used a separate scan deliberately taken for noise measurement to extract noise. This separate scan could be used as a baseline for noise level estimation to generate background noise. The noise data was read using ReconFrame by Gyrotools (www.gyrotools.com/reconframe). The standard deviation of the raw noise data was calculated and used to generate Gaussian noise with a similar spread. However, this method proved to be insufficient as the standard deviation differed too strongly from the standard deviation of the original image's noise. The mean standard deviation of the raw noise data was around 2.3, while the noise of the background region was between 6,000-9,000. Thus, a better method that addressed these problems was needed.

The next method that was explored worked by using k-space from the complex coil images to generate noise. The complex coil image data was converted into k-space data and normalised for the change with a factor of $\frac{1}{\sqrt{n}}$, with n referring to the number of data points. The last slice of data in the first dimension with size 292 was taken and a ring was selected. The outer circle had a radius of 117 pixels while the inner circle had a radius of 90 pixels. The ring itself thus had a width of 27 pixels. Within that ring, the standard deviation was calculated. As the data stemmed from the last slice of k-space and was far from the centre of k-space, it was assumed that the data in this ring would mainly consist of noise. Thus, the standard deviation was calculated and with this number as the noise level and Gaussian noise was generated using $\mu = 0$. This newly generated noise was then multiplied with an inverse version of the dilated brain mask, after which it was combined with the defaced image. An example of the k-space slice and corresponding ring used to calculate the standard deviation for the noise can be found in Figure 4.3. The noise level was calculated for each coil of the complex coil image. The noise was generated separately for the real and imaginary part of the data.

To create a defaced image with the same noise levels as an original scan, the generated Gaussian noise was compared to the existing noise in the original image. For this, a $50 \times 50 \times 50$ voxel cube was selected in the background of the original image and the standard deviation was calculated. This sample was compared to a sample of the same location for the defaced image. A depiction of this box of the background can be found in Figure 4.4.

To ensure a smooth transition between the added Gaussian noise and the defaced image, alpha blending was used to create a composite. A binary inversion of the defaced image was created whose edges were smoothed using a Gaussian filter. This smoothed inversion was then merged with the smoothed defaced image to create one defaced MRI scan.

4.2.4 Publishing AHEAD

The final goal of the defacing was to be able to publish the AHEAD data online and make it available to a wider public. Because of the large size of MRI data, a hosting facility was needed. It was decided to use Dataverse in association with Data Archiving and Networked Services (DANS). DANS is a [Technology Hotel](#), which means that they have knowledge, expertise, and data storing facilities, that the Amsterdam UMC might not have on their own. [Dataverse](#) was originally established by Harvard University, Cambridge, MA, USA and has been widely used to host scientific data (Holmes et al., 2015; Tona et al., 2019). The intent for this task was to release the data under a Creative Commons Attribution 4.0 International Public License, which requires attribution when using the data. Currently, all data is accompanied by a subject ID. In order to assure anonymity, these subject IDs will be randomly shuffled twice after which the key will be deleted. This is to ensure that nothing can trace back a subject to an individual person.

4.3 Quantitative MRI Reconstruction using the RIM

As mentioned before, the data used for the sharpness method consisted of MRI scans reconstructed using the Recurrent Inference Machine (RIM). Contrary to a usual MRI scan, the quantitative MRI used in this work did

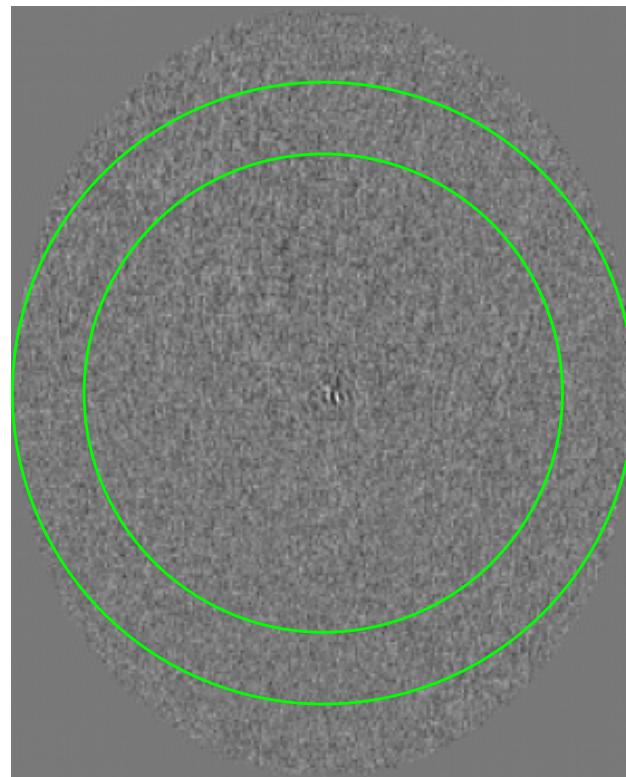


Figure 4.3: Image depicting the ring used to calculate the standard deviation (STD) used to generate background noise. The values in the area between the two green circles is used.

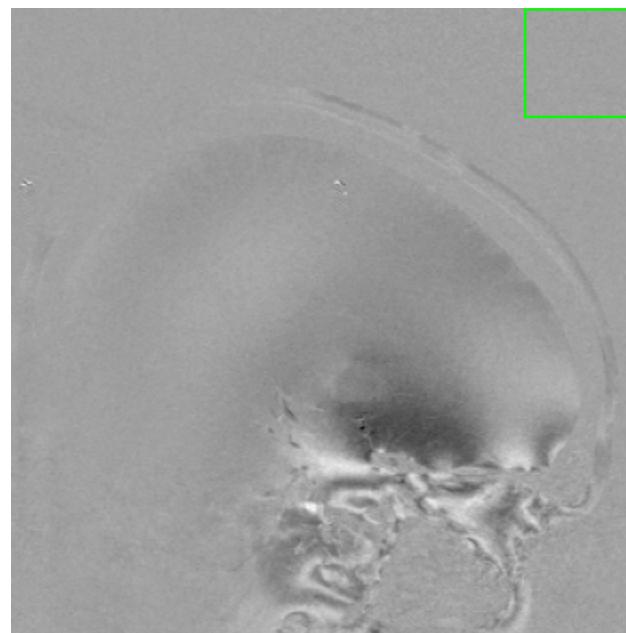


Figure 4.4: 2D single-slice depiction of the location of the cube that is used to compare the noise level in the original image and the defaced image.

not mix the echos in one k-space. This method typically acquires a separate image for each echo time. However, fully sampling k-space is infeasible due to its long acquisition time. Therefore, k-space is usually undersampled to reduce scan time. Unfortunately, this results in heavy aliasing, i.e. overlapping of structures (Westbrook et al., 2011), in the reconstructed images. Because of this, the undersampled k-space needs to be reconstructed using advanced methods. While multiple reconstruction methods exist, as discussed in Section 3.2, the RIM was used for its merits in reconstruction quality and robustness against overfitting.

At the basis of MRI reconstruction using the RIM lies the forward model (Lønning et al., 2019). MRI reconstruction for parameter mapping is solving an inverse problem where k-space measurements need to be mapped to relaxation parameters. The forward model maps relaxation parameters to an image of echo time TE_t and further the k-space signals. Let $\mathbf{x} \in C^n$ be the true image signal and let $\mathbf{y}_t \in C^m$, where $m \ll n$, be the set of undersampled frequency signals measured in k-space by one of the scanner's c receiver coils. The true image x_t can be described by

$$x_t = M e^{TE_t(R_2^* + B_0 i)} \quad (4.3)$$

where R_2^* refers to the parameter map of R_2^* and B_0 represents the off-resonance of the static magnetic field (Zhang et al., 2021). The undersampled measurements can then be described in terms of the true image x_t by

$$y_{t,c} = P \mathcal{F} S_c x_t + n_c \quad (4.4)$$

where S_c represents the coil sensitivity map, \mathcal{F} denotes the Fourier Transform used to project the image from frequency domain to image domain, P is the sub-sampling mask. The sensitivity map is a diagonal matrix where every pixel was scaled by a complex number, all according to the spatial sensitivity of the c th receiver coil. The sub-sampling mask reduces the dimensionality by discarding a fraction of k-space values, thereby enabling the acceleration at the time of scanning. Measurements can be subjected to additive Gaussian noise denoted by

$$n_c \sim \mathcal{N}(0, I\sigma^2) + i\mathcal{N}(0, I\sigma^2) \quad (4.5)$$

Combining Equation 4.3 with Equation 4.4 yields the following unified forward model (Zhang et al., 2021):

$$y_{t,c} = P \mathcal{F} S_c M e^{TE_t(R_2^* + B_0 i)} + n_c \quad (4.6)$$

4.3.1 RIM update equations

The main goal of accelerated MRI reconstruction is to find an inverse transformation for the given forward model in Equation 4.3 (Lønning et al., 2019). It is common to achieve this by optimising the maximum a posteriori (MAP) estimator given by

$$\Phi_{MAP} = \arg \max_{\Phi} \{\log p(\mathbf{y}|\Phi) + \log p(\Phi)\} \quad (4.7)$$

Here, Φ refers to the following:

$$\Phi = \{R_2^*, M, B_0\} \quad (4.8)$$

where R_2^* once again refers to the parameter map of R_2^* , M refers to the net transverse magnetisation and B_0 is the off-resonance of the static magnetic field. Effectively, this is the maximisation of the sum of the log-likelihood and log-prior distributions of Φ and \mathbf{y} . This MAP estimate is often reformatted as optimising the regularised problem. This problem is formalised as follows:

$$\arg \min_{\Phi} \left\{ \sum_{t=1}^T \sum_{c=1}^C d(\mathbf{y}_{t,c}, P \mathcal{F} S_c x_t) + \lambda R(\Phi) \right\} \quad (4.9)$$

where d ensures data consistency between the reconstruction and measurements, and R is a regulariser, with λ as a regularisation factor. The regularisation constrains the solution space further and prevents overfitting. It does so by incorporating prior knowledge about the solution.

Under the assumption that the data is identically and independently distributed (i.i.d.), the negative log-likelihood of this problem then becomes:

$$\log p(\mathbf{y}|\Phi) = \frac{1}{\sigma^2} \sum_{t=1}^T \sum_{c=1}^C \|P \mathcal{F} S_c M e^{-TE_t(R_2^* - \hat{B}_0)} + n_c - \mathbf{y}_{t,c}\|_2 \quad (4.10)$$

The aim of the RIM is to optimise Equation 4.9 (Lønning et al., 2019). It does so by learning an iterative scheme over τ recurrent timesteps. As per the recurrent nature of the RIM, each timestep takes as input the information about the current state of the reconstruction which yields an incremental step $\Delta\Phi_\tau$ with parameter

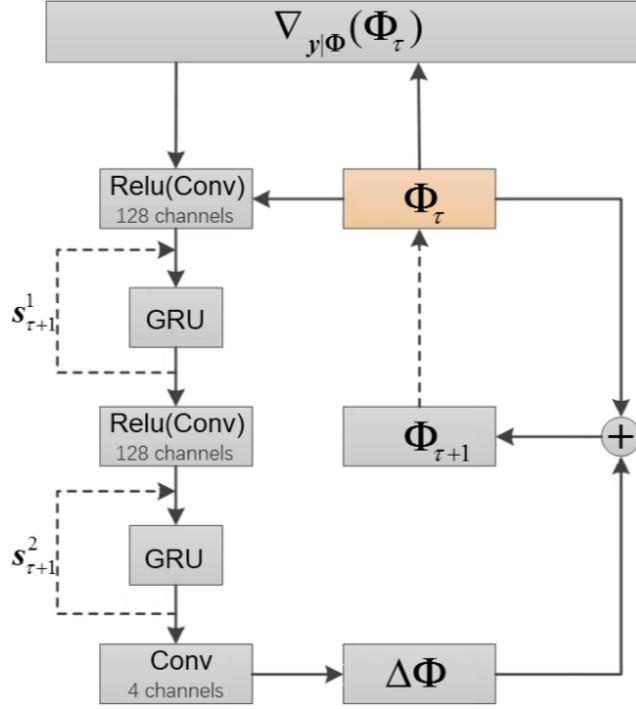


Figure 4.5: An illustration of the network architecture of the RIM. Figure created by [Zhang et al. \(2021\)](#).

space as output. One such input is the gradient of the log-likelihood as presented in Equation 4.10, for timestep τ .

Now, let the an update of the RIM network be denoted by h in such a way that each pass through h yields the next incremental update to $\Delta\Phi_\tau$. The RIM update equations are then as follows:

$$\mathbf{s}_0 = 0 \quad \Phi_0 = \sum_{t=1}^T \sum_{c=1}^C S_c^H \mathcal{F}^{-1} P^T \mathbf{y}_{t,c} \quad (4.11)$$

$$\mathbf{s}_{\tau+1} = g(\nabla_{y|\Phi_\tau}, \Phi_\tau, \mathbf{s}_\tau) \quad \Phi_{\tau+1} = \Phi_\tau + h(\nabla_{y|\Phi_\tau}, \Phi_\tau, \mathbf{s}_{\tau+1}) \quad (4.12)$$

for $0 \leq \tau < t$ ([Lønning et al., 2019](#)). Here, g represents the part of the network that is responsible for producing the next internal state $\mathbf{s}_{\tau+1}$. The RIM requires this next internal state to keep track of iterations and modify its behaviour based on the state of the inference procedure.

4.3.2 Update function

To implement the RIM's update function h , a sequence of alternating convolutional layers and gated recurrent unit (GRU) cells was used. After the first two convolutional layers, a ReLU activation function was inserted before feature maps were passed to the GRUs. The GRUs ensured the internal state \mathbf{s} was preserved and were based on the RNN Encoder-Decoder network as presented by [Cho et al. \(2014\)](#). Due to the Encoder-Decoder nature of the GRU, there are two internal states $\mathbf{s} = \{\mathbf{s}_1, \mathbf{s}_2\}$. In the update functions, this is simply denoted by \mathbf{s} (see Equations 4.11 and 4.12). The input of the first convolutional layer is produced by concatenating the external state Φ_τ with the log-likelihood gradient $\nabla_{y|\Phi_\tau}$ along the channel dimension. This results in four input channels as the complex components are stored in separate channels. A kernel size of 5×5 was used for the first layer, while all other convolutional layers have a kernel size of 3×3 . To ensure the original image size is kept throughout the network, all convolutions are padded.

The number of feature maps produced by the convolutional layers and the number of features in the GRU are considered a hyperparameter that is kept the same throughout all internal layers. This hyperparameter is denoted by f . It should be noted that the GRU shares the weights of its cells across image pixels. However, these weights differ across feature maps and allow the network to reconstruct images of any given resolution.

An illustration of the RIM can be found in Figure 4.5.

4.3.3 Loss function

The RIM uses the structural similarity (SSIM) as a loss function ([Zhang et al., 2021](#)). The estimate Φ_τ is evaluated against the true image x_t for each timestep. The total loss is the weighted sum of the SSIM for all

timesteps, as denoted by

$$\mathcal{L}(\Phi_\tau) = \frac{1}{nt} \sum_{\tau=1}^t \|\Phi_\tau - \Phi_{ref}\|_2^2 \quad (4.13)$$

where n is the number of pixels in the image and t is the number of timesteps during training. w_τ is a parameter that determines how much emphasis is put on image quality during reconstruction τ relative to the other $t - 1$ estimates.

In this work, the focus will mainly lie on R_2^* maps. This choice was made because, when applying the forward model of the RIM, the use of R_2^* maps is more stable. Through the forward model of signal relaxation, small turbulence around a small valued T_2^* can lead to unstable output. Due to the nature of the inversion for R_2^* , this problem is avoided.

4.4 Sharpness estimation along tissue boundaries

Measuring sharpness in MR images is not a trivial task. In this work, the sharpness measure developed by [Bazin et al. \(2020\)](#) was used. Originally developed for determining the sharpness of T_1 motion-corrected images, this procedure was re-purposed for determining sharpness in RIM reconstructed R_2^* images for different acceleration factors.

Sharpness was described as the edge width of a specific region of interest (ROI) where sharpness was defined along the transition boundary to the surrounding tissue ([Bazin et al., 2020](#)). The intensity transition was modelled by using a sigmoidal function. The actual sharpness estimation over the edge of the ROI was fitted to signal values. This allowed for an accurate estimation of the edge width and the corresponding uncertainty, taking into account the amount of noise in the data. The eventual sharpness was measured in physical distance, which made the estimation invariant to scaling. To control for variability in tissue signal, a clustering approach was utilised to select sub-regions of homogeneous tissue over which to calculate sharpness. Calculations were also restricted to clusters where sufficient contrast was present with the surrounding tissue. This was assessed by the model's goodness of fit.

4.4.1 Distance function

To measure distance, the point spread function (PSF) was used. Usually, in MRI, the PSF is considered to be an isotropic sinc(\cdot) function (see Equation 4.14). However, when using the MP2RAGE-sequence with Cartesian sequential partition-encoding (as was the case here), this PSF is assumed to be directional because of the acquisition of T_1 -weighted images. Furthermore, patient motion can cause a detrimental effect on image quality, together with ghosting and other artefacts. Additionally, sharpness can also be influenced by a potential gradual change of tissue. All these effects should be taken into account when creating a sharpness measure. Due to the difficulty of this task, [Bazin et al. \(2020\)](#) took a more practical approach.

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \quad (4.14)$$

First, the PSF was locally estimated and was assumed to be Gaussian and isotropic. The PSF is defined as follows:

$$PSF = \mathcal{G}(\sigma^2) \quad (4.15)$$

where σ denotes the standard deviation. It was assumed motion only affected σ and nothing else. Lastly, it was assumed that a tissue boundary was an edge function, which, after being exposed to a Gaussian PSF, yielded a sigmoid function.

Its profile can then be expressed in the form of a so-called error function ($\text{erf}(\cdot)$), which is a cumulative Gaussian function. The measured signal as a function of distance $S(d)$ is then as follows:

$$S(d) = S_0 + \frac{h}{2} \cdot \text{erf} \left(\frac{d - d_c}{\sqrt{2}\sigma} \right) + \mathcal{N}(0, \tau^2) \quad (4.16)$$

Here, S_0 is the signal offset, h is a scaling factor that denotes the difference in contrast over the edge width. d_c is the spatial offset to control for a possible local segmentation offset with respect to the current model. It is assumed that the image contains some additive noise \mathcal{N} with variance τ^2 . Sharpness is then inferred from the full width at half maximum, which is denoted as $\text{FWHM} = 2.36 \cdot \sigma$.

4.4.2 Clustering

An overview of the clustering approach is visualised in Figure 4.6. Because the internal width of each ROI differs, the clustering was performed only in the external region (see Figure 4.6b). The process consisted of two stages. The first stage was a k-means clustering of z-transformed and concatenated spatial coordinates and signal values. Neighbouring regions with similar signal intensity were grouped together in this initial clustering stage. This way, continuous regions with comparable signal intensity are grouped together (see Figure 4.6c). In order to suppress noise, intensities were smoothed using a Gaussian filter with a standard deviation of three voxels. In the clustering stage, tissue was divided into four classes by type. The second clustering stage was used to ensure spatially coherent clusters, as these are not guaranteed to occur when the clustering is performed with signal values as a feature. This was done by clustering only the spatial coordinates within each of the previously obtained clusters (see Figure 4.6d). To eventually obtain regions in which signal profiles were derived and curve fitting could be performed, all voxels in the external and internal part of the ROI were assigned to a cluster (see Figure 4.6e). This was done by following a nearest-neighbour rule.

To estimate sharpness, spatially coherent regions were selected for curve fitting. These spatially coherent regions were sampled from either side of the ROI boundary, ranging from 3 mm internal to 4 mm external. This external part could not overlap with multiple anatomical structures. The average cluster size was set to 500 voxels, which corresponds to a sphere with a radius of 5 voxels. As a consequence, the number of clusters per ROI can vary significantly depending on the size of the ROI. Within each cluster, the normal distance d for each voxel to the boundary was calculated. The sigmoid function from Equation 4.16 was then fitted to these normal distances and used to calculate the full width at half maximum (FWHM).

K-means clustering works as follows. Given a set of data points x_1, x_2, \dots, x_N and a corresponding set of binary indicator variables $r_{nk} \in \{0, 1\}$, where $k = 1 \dots K$ describes to which cluster k this datapoint belongs (Bishop, 2006). Then, the following objective function can be derived:

$$J = \sum_{n=1}^N \sum_{k=1}^K r_{nk} \|x_n - \mu_k\|^2 \quad (4.17)$$

where μ_k represents the centre of the cluster k . The objective is then to find values for r_{nk} and μ_k such that J is minimised. This, in essence, is minimising the within cluster sum of squares (WCSS).

After finding the clusters and assigning tissues to them, profiles following the desired sigmoidal shape were marked as valid, while others were marked as invalid. This validation check showed whether the method was able to determine accurate sharpness.

4.4.3 Coregistration

The sharpness method explained above made use of levelmaps to segment sub-cortical regions for sharpness estimation. These levelmaps allowed for sub-voxel segmentation by using values above zero to indicate voxels outside the ROI while values below zero to indicate voxels within the ROI. An example of a levelmap can be found in Appendix A, Figure A.1. These levelmaps were thus able to follow a more precise boundary than binary maps would. However, in the initial data creation, these levelmaps were only created for a limited number of subjects. Furthermore, the creation of the $R2^*$ maps involved training with a certain number of subjects, limiting the number of subjects for the eventual output of the reconstruction network even further. The intersection between these two limited sets of subjects resulted in a total of three subjects being available for this sharpness estimation. As this number was deemed insufficient, we set out to acquire additional segmentations.

Binary segmentation maps were available for more subjects, but these were generated based on a different scan than the $R2^*$ scans. To ensure the usability of these scans, two steps needed to be taken. The binary maps needed to be transformed into levelmaps to retain the sub-voxel resolution, and these maps needed to be coregistered in order to align them with the scans used for reconstruction. Luckily, tools exist for both tasks.

Nighres (Huntenburg et al., 2018) offers a module that creates a levelmap from a binary mask and FLIRT (FMRIB's Linear Image Registration Tool) (Jenkinson and Smith, 2001; Jenkinson et al., 2002) is a tool that allows for coregistration of MRI volumes. This FLIRT tool works across modalities, so it is possible to register a $R1$ image with a $R2^*$ image. The process of coregistration in essence aims at finding a transformation T such that the input image I_i most resembles the reference image I_r . This is done using a cost function $C(I_1, I_2)$. The problem is formulated as follows:

$$T^* = \arg \min_{T \in S_T} C(I_i, I_r \circ T) \quad (4.18)$$

where S_T is the search space with transformation matrices T that are valid for this problem. In this work, the Correlation Ration, a cost function proven to function well for between-modality registration, was used with 12 degrees of freedom (DOF).

The coregistration procedure worked as follows: the $R1$ scan from scan-moment 1 was loaded in and coregistered with the $R1$ scan from scan-moment 2, the scan-moment from which the $R2^*$ maps were created. This

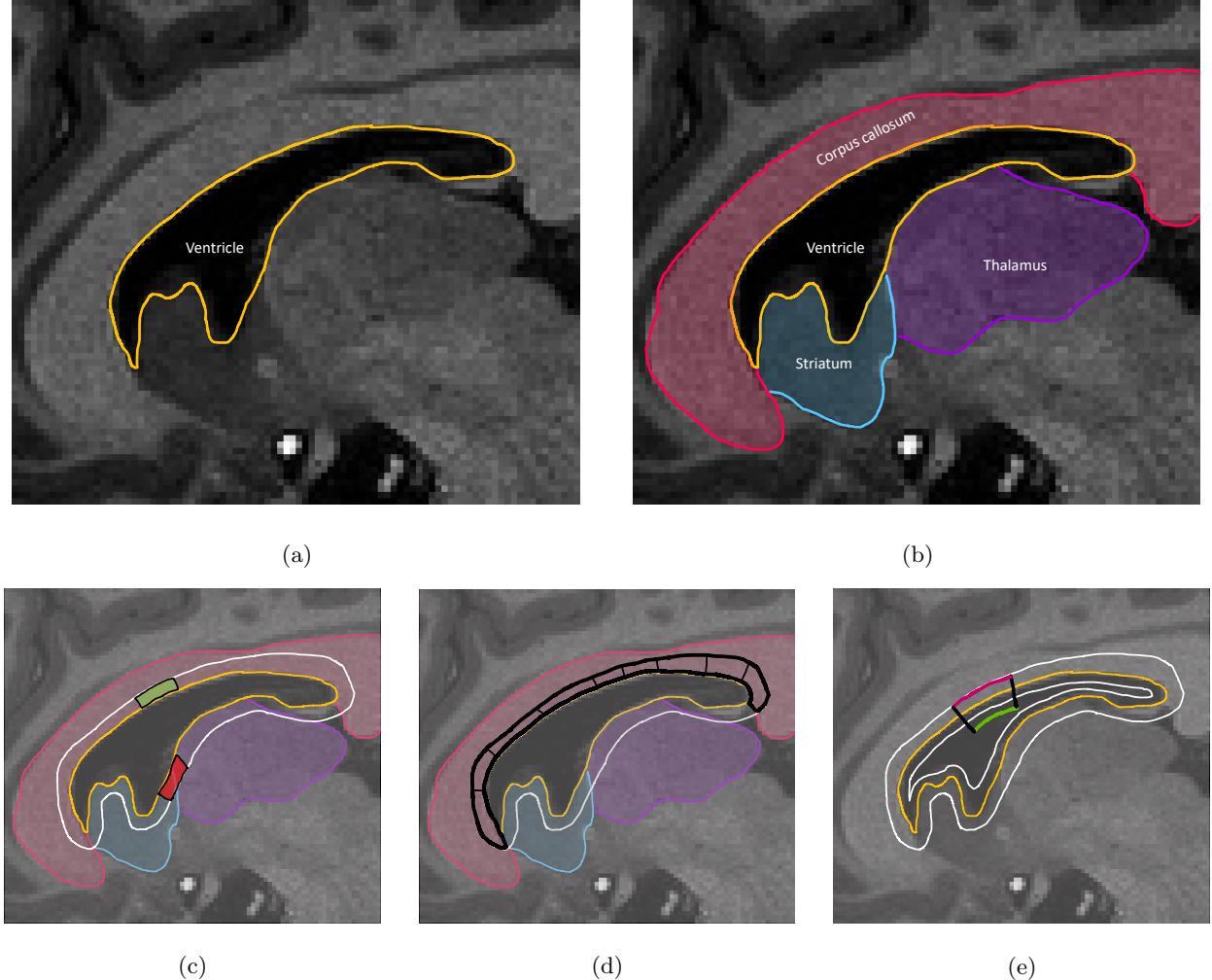


Figure 4.6: Illustration of the clustering approach. (a) Example of an ROI. Here, the ventricle is outlined. (b) The chosen ROI might border multiple other anatomical structures. (c) Borders after the expansion of the ROI. Part one of the clustering approach determines clusters that are homogeneous in tissue. The green (top) area indicates a homogeneous region while the red (bottom) area marks an inhomogeneous region, as it overlaps multiple regions. (d) A two-step hierarchical k-means clustering is performed in order to locate spatially coherent clusters within areas of homogeneous intensity. (e) Clusters are expanded to include the ROI to compute intensity profiles.

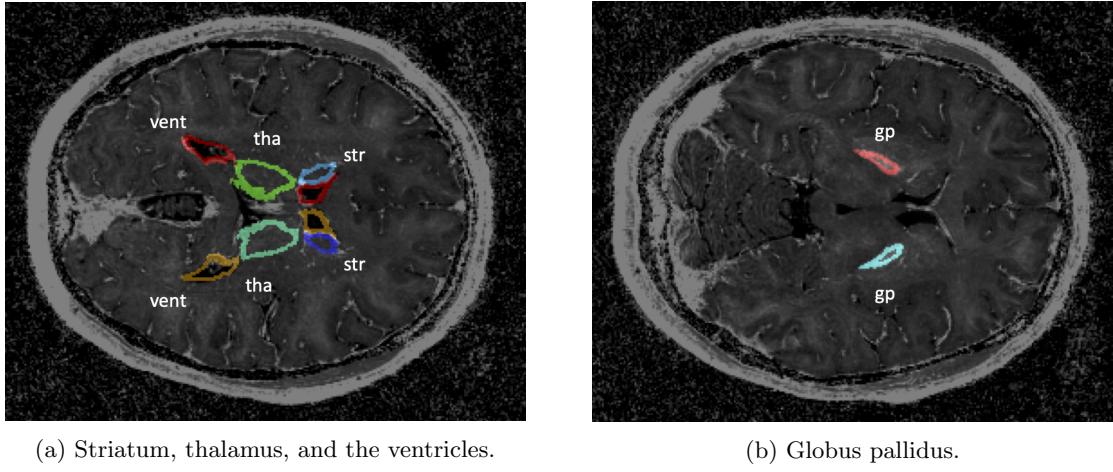


Figure 4.7: Outline of the used ROIs in the sharpness estimation. Pictured are the striatum (blue, light blue), thalamus (green, lime green), and the ventricle (red, orange) (4.7a), and the globus pallidus (4.7b) in both the left and right hemisphere.

coregistration resulted in a transformation matrix T . The binary mask was then separated into a total of 32 regions, after which each region was used separately to create a levelmap using Nighres. Afterwards, the transformation matrix T was applied to this newly created levelmap using a built-in function of FLIRT. Finally, the levelmap was cropped to maintain the same dimensions as the R_2^* map. The result is now a fully coregistered levelmap of each region and can be used in the sharpness estimation. For convenience, the full binary map was also coregistered. To control for interpolation caused by coregistration, the binary map was converted to integers to control for this.

4.4.4 Subjects and ROIs

A total of nine subjects were used for sharpness estimation. Four regions were explored: the globus pallidus, the striatum, the thalamus, and the ventricles (excluding the fourth ventricle). The sharpness was estimated separately for both the left and the right hemispheres. An outline of the structures used can be found in Figure 4.7.

Chapter 5

Results

5.1 Defacing

5.1.1 Defacing using Freesurfer

Because the T_1 images that were available contained a noisy background, a first step was to determine whether these T_1 images could be used or whether another contrast would be required. To find the optimal image to perform the defacing on, the Freesurfer defacing was applied to two types of scans: a T_1 motion-corrected scan and an image of the first echo of the second inversion (Caan et al., 2019). Figure 5.1 depicts the T_1 image in the left column and the first echo of the second inversion in the right column. The defacing appears to have been unsuccessful for the second inversion while functioning well for the T_1 image, being able to detect the face and successfully remove it. The Freesurfer defacing removed the face and part of the background area, leaving the image with an uneven gap and distribution of background noise. However, it was decided to solve this issue once it was established that the Freesurfer defacing worked sufficiently for all subjects. Because the T_1 corrected images produced the best result, this type of scan was used in removing facial structures from other subjects as well.

However, when the defacing of multiple subjects was explored, results showed that the Freesurfer defacing was not able to generalise very well. Results of this can be found in Figure 5.2 where the defacing of three different subjects is compared. The Freesurfer defacing software did not always remove the entire face, leaving facial features intact. In these cases, the software removed the inner part of the face but left an outline of the face untouched. As this would allow facial reconstruction, the procedure did not meet the criterion for anonymity.

5.1.2 Defacing using dilation

Because dilation is very versatile, the first step in developing a defacing algorithm was to determine which kernel size and how many iterations were needed to optimise facial removal. After initial testing, it was determined that a 3D $5 \times 5 \times 5$ kernel would be sufficient if applied for a certain number of iterations. An illustration of the dilation with a different number of iterations is displayed in Figure 5.3. Dilation using only five iterations has shown to exclude any facial regions, but also exclude the skull, which should be included. Dilation using 20 iterations included too much of the eyes, which could allow the identification of a specific person. Thus, the choice was between 10 and 15 iterations. In the end, dilation with 10 iterations was used because this included the skull, but the least facial information and was therefore deemed the safer option.

To explore whether this number worked for other subjects as well, the algorithm was applied to three different subjects. Results can be found in Figure 5.2. The dilation removed the face in all three subjects, but left the skull intact, showing that the chosen kernel size and number of iterations was able to generalise to multiple subjects.

Once it was decided to use a $5 \times 5 \times 5$ kernel for 10 iterations, the defacing algorithm was applied to the data. Figure 5.5 shows the results for the defacing by dilation of the brain mask applied to a T_1 motion-corrected scan. This result is shown for different planes to get a better impression of how the dilation affects the scan in different directions. Noise is an important part of the data and has not yet been added to the background. The algorithm removed facial features in all directions, leaving no hidden facial features behind. Further, it did not remove the skull, except at the area around the neck.

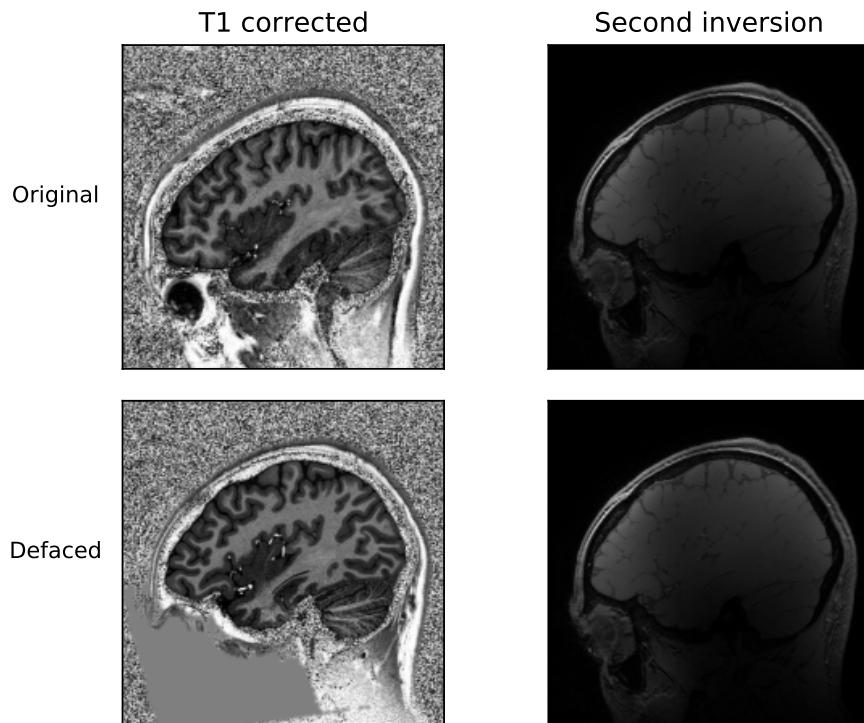


Figure 5.1: Results of defacing using the Freesurfer defacing software for a single subject. Defacing was performed on the original images.

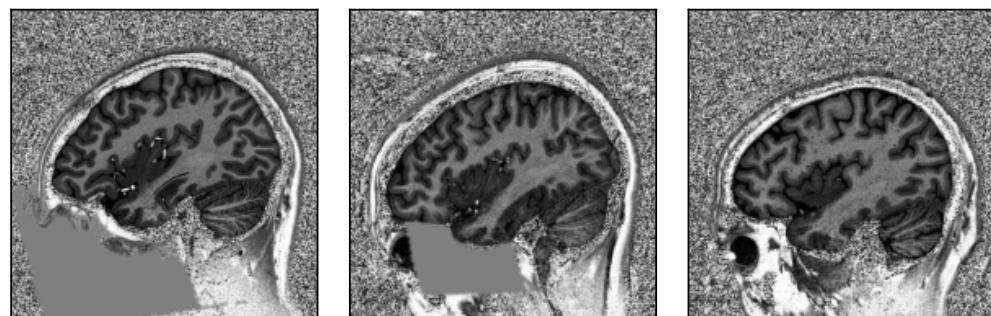


Figure 5.2: Freesurfer defacing applied to three different subjects.

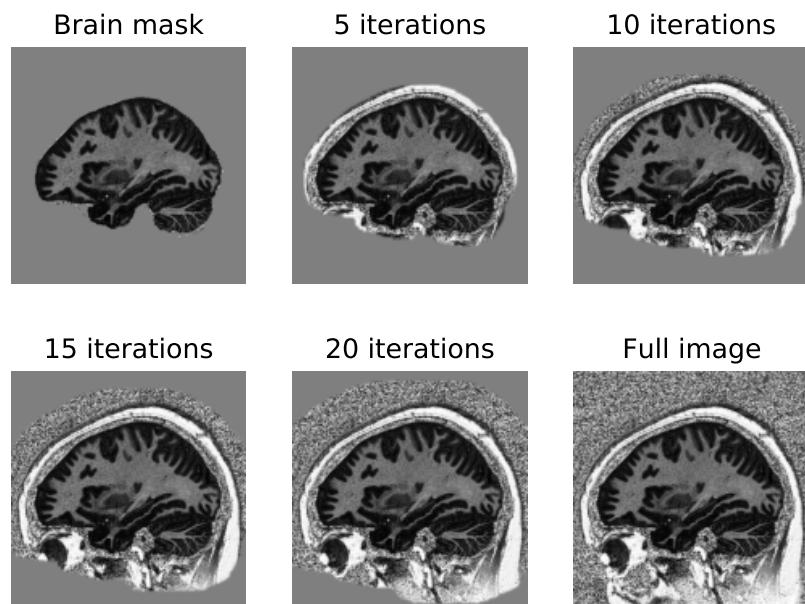


Figure 5.3: Defacing using dilation of the brain mask for different numbers of iterations. A kernel of $5 \times 5 \times 5$ was used for all iterations.

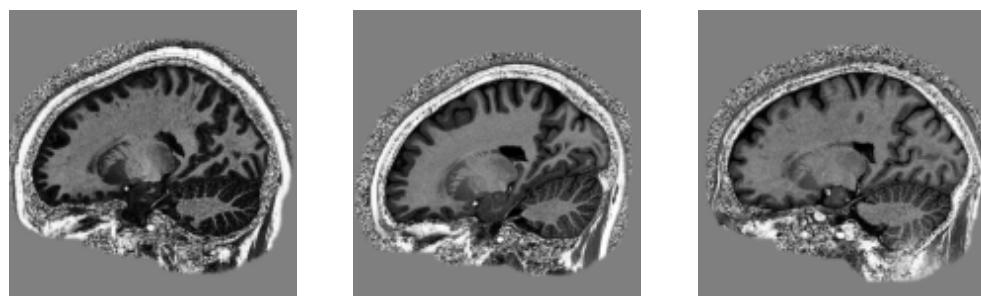


Figure 5.4: Defacing by dilation applied to three different subjects.

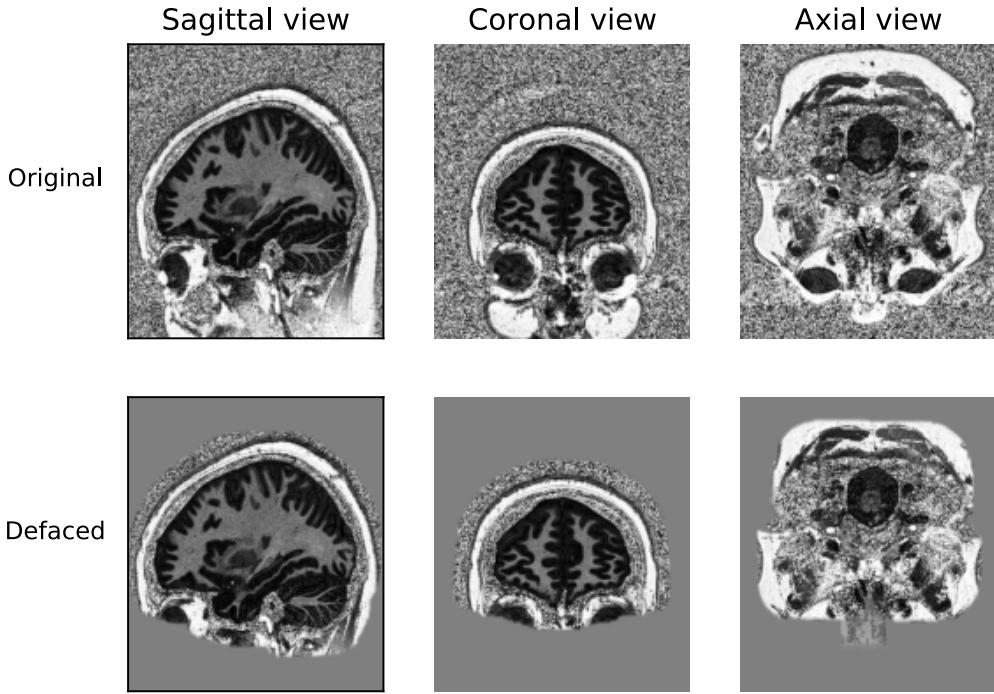


Figure 5.5: Defacing results using the dilation algorithm shown for different planes in one subject.

5.1.3 Adding Gaussian background noise

Figure 5.6 shows the difference in percent between the noise in the box of the original scan and the defaced scan with added noise for all subjects combined. The blue line depicts the mean difference of all coils of the real and imaginary numbers combined. This graph demonstrates a large difference in the noise estimation, averaging at 47.5%. A t-test showed that this difference is significant for both the real and imaginary channels, with $p < 0.001$ for both. A second t-test showed no significant difference in the noise of the real and the imaginary channels ($p = 0.98$). The estimation of the noise appears to follow a similar pattern for all echo times, with the estimation being more similar to the original in coils 2-5 and coils 23-27. The average overestimation also appears to be stable across echo times, with the second echo time showing the highest overestimation at 53.7% and the fourth echo time showing the lowest overestimation at 42.8%.

To investigate whether there was a smooth transition between defaced image and the added noise, a single-coil image was plotted with a different colour range to visualise the boundary. This image can be found in Figure 5.7. In the image, the overestimation of noise can be seen in the darker speckles on the background, as opposed to the lighter speckles in the original image. A small boundary can also be seen along the back of the skull, proving that the transition is not completely smooth. However, the boundary does not appear to cause a major disruption.

5.2 Sharpness estimation along tissue boundaries

The sharpness estimation was initially developed for T_1 -weighted images (Bazin et al., 2020) but it was not contrast-dependent and was used on R_2^* maps in this work. These maps did not always contain enough contrast for the sharpness method to provide an estimation. When a sharpness estimation is a good indicator for image quality, the method would show an increase in FWHM for each acceleration factor. In order to be able to draw a conclusion, results will first be explored on a subject and brain region level, then on a region-only level, after which the view will shift to an overall subject level. FWHM is an indicator for the smoothness of the image, which will be interpreted as sharpness in this work. Because of that, a lower FWHM means a smoother and therefore sharper image.

Figure 5.8 shows the estimated FWHM for one subject for the globus pallidus, the striatum, the thalamus, and the ventricles. For each region, the ground truth was the same for all acceleration factors which makes it possible to compare acceleration factors for reconstructed images. The sharpness for the ground truth and the reconstructed image was quite similar for the acceleration factor of 3, but the FWHM diverged with a higher acceleration factor. This is especially visible in the globus pallidus (gpl and gpr). There is no FWHM displayed for the thalamus (thal and thar), as the sharpness estimation did not have enough contrast with surrounding

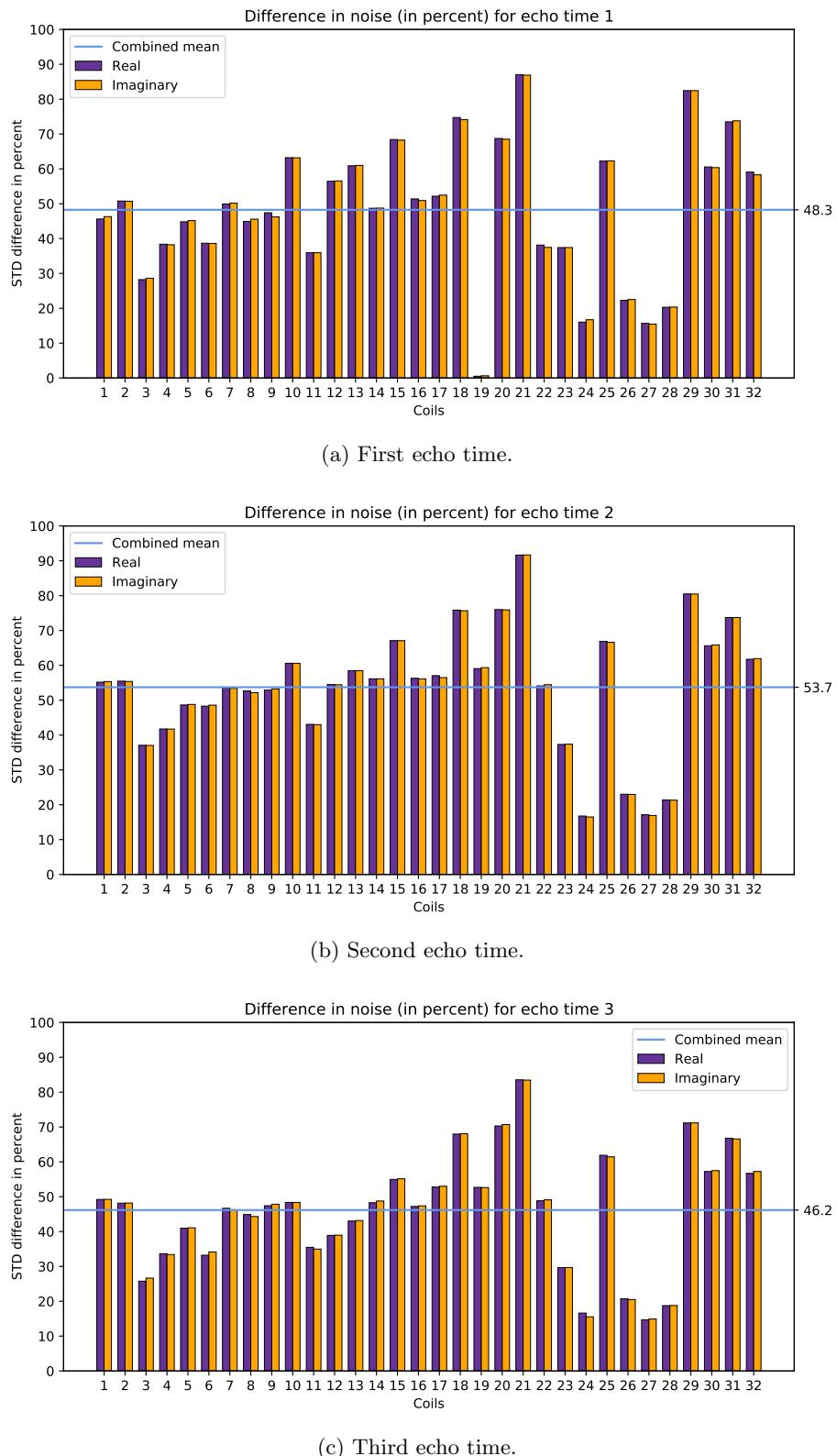
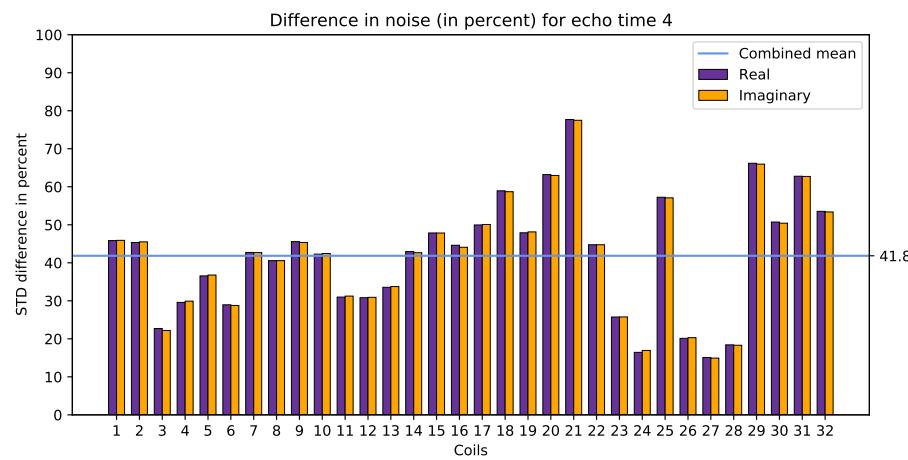


Figure 5.6: The percentage difference in noise from the original image compared to the newly generated noise of the defaced image.



(d) Fourth echo time.

Figure 5.6: The percentage difference in noise from the original image compared to the newly generated noise of the defaced image.

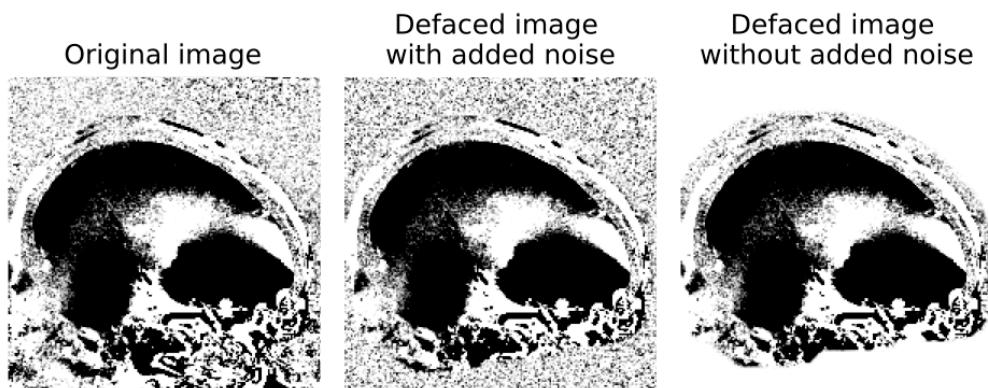


Figure 5.7: Comparison of the original image and the defaced scan.

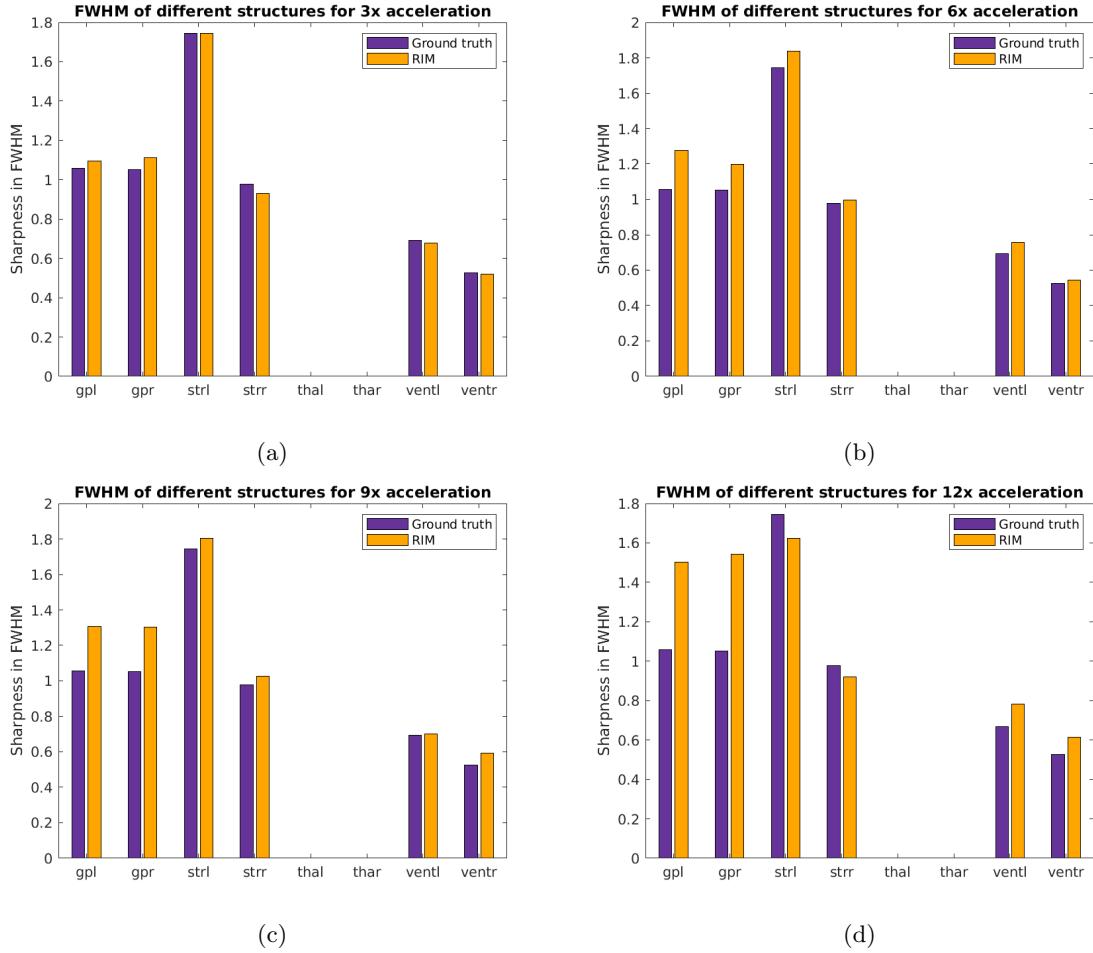


Figure 5.8: The FWHM of one subject for different acceleration factors for both ground truth and accelerated image. Ground truth is kept at the same level for better comparison.

regions to give an estimate.

The need for contrast in the estimation is especially visible in the generated tissue boundary profiles, visible for the ground truth and acceleration factor 12 in Figure 5.9. These profiles were generated for the same subject as the results depicted in Figure 5.8. Figure 5.9a shows that the clustering method is able to find a sigmoidal shape for the globus pallidus in the right hemisphere for all but one of the available profiles. The figure below (Figure 5.9c) shows the profiles for the thalamus. Here, the method was not able to find any valid profiles, with the lines all showing a shape deviating from the desired sigmoidal shape. The tissue boundary could either not be found or was deemed invalid, classifying it as a gradual change instead of a clear boundary, not allowing for sharpness estimation. Profiles for all acceleration factors can be found in Appendix B, Figures B.1 and B.2.

To allow for a comparison between acceleration factors, the sharpness estimation of the ground truth must always return the same value. Because the selection of subclusters over which the sharpness was estimated had a random component, this was different for each iteration of the estimation. Thus, this subclustering method was fixed by saving the chosen subclusters of the first run of the algorithm. For the sake of continuity, all initial runs were performed with an acceleration factor of 3. Initial inspection of the data showed that this ground truth value still differed between runs. This is likely due to the fact that these regions were too low-contrast. As a consequence, the algorithm labelled some of these subclusters as invalid for higher acceleration factors. To allow for a reliable comparison, these outlier results were removed from the dataset. The data over which statistical analysis was performed showed equal ground truth calculations for all regions.

5.2.1 Trend analysis

To investigate whether an overall upwards trend of FWHM was visible in the sharpness estimation, the mean FWHM per region for all subjects was plotted (see Figure 5.10). In most brain regions, there is only a small difference between the ground truth data and the 3 times accelerated data. At acceleration factors of 6 or higher an increase in FWHM begins to appear. Overall, the plot still shows an upwards trend in FWHM for regions such as the globus pallidus, but less of a trend for areas such as the ventricles.

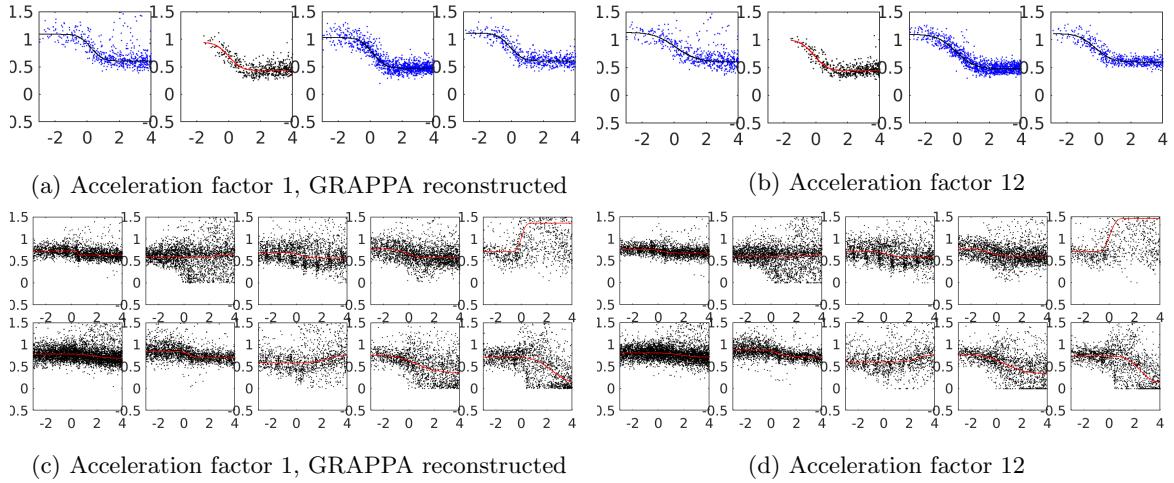


Figure 5.9: Profiles of the sharpness estimation showing whether a boundary is valid and contrasting enough to be included in the sharpness estimation.

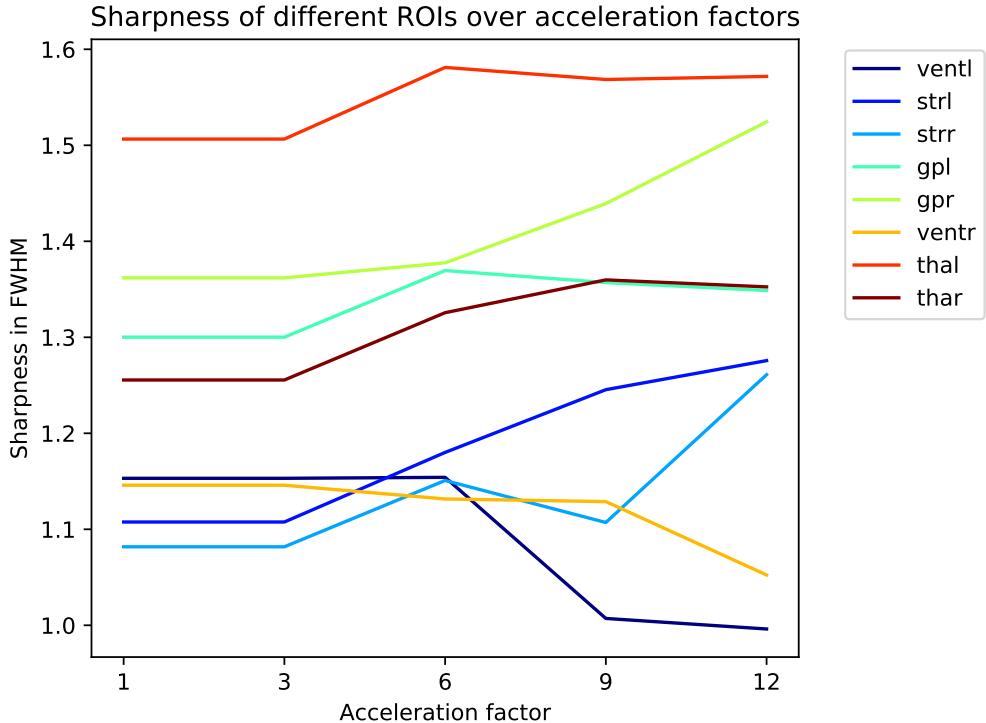


Figure 5.10: Mean FWHM for all subjects per region for all acceleration factors. Gp denotes the globus pallidus, str the striatum, tha the thalamus, and vent denotes the ventricles. All regions are displayed for both the left (-l) and right (-r) hemispheres.

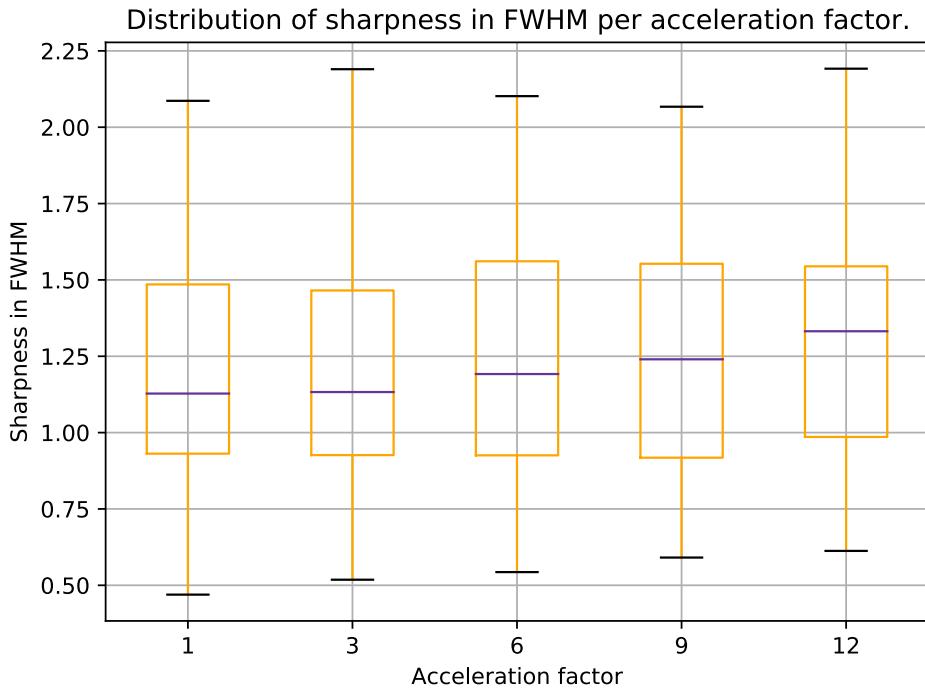


Figure 5.11: Depiction of the distribution of sharpness in FWHM for all regions per acceleration factor.

The sharpness for images with acceleration factors 3 and 12 was compared. Figure 5.11 depicts a boxplot with the distribution of the FWHM for all acceleration factors, giving insight into the spread of the data. The medians of the ground truth (acceleration factor of 1) closely matched the median of the FWHM for acceleration factor 3, which was also visible in Figure 5.10. The boxplot further shows that the median increases with each subsequently higher acceleration factor, indicating a positive correlation between acceleration factor and median. However, the figure also shows that the data is widely spread, with the first and fourth quartile being quite similar, making it difficult to confirm a trend.

In order to determine a trend, a t-test was performed on the FWHM from 3 times accelerated and 12 times accelerated MR images. While not all regions were high in contrast, it was decided to nonetheless include all regions in the trend analysis to ensure the availability of enough data points. Although an initial histogram indicated that the data might not be normally distributed (see Appendix B, Figure B.3), a more thorough normality testing proved the data was normally distributed with a p-value of 0.836 for the 3 times accelerated data and a p-value of 0.439 for the 12 times accelerated data. A Levene test showed that the data had equal variances ($p = 0.588$), so a t-test could be performed.

Because the results contained multiple samples of the same subject, a paired t-test was carried out. The data contained a separate measure for both the left and the right hemispheres. However, a paired t-test assumes independent samples within each measure. To comply with this assumption, the FWHM from each hemisphere for each ROI was averaged. A paired t-test showed a significant difference between the FWHM for 3 times accelerated images and the FWHM for 12 times accelerated images with a p-value of 0.029.

Chapter 6

Discussion

6.1 Defacing

6.1.1 Defacing method

The Freesurfer defacing method proved to be insufficient for removing all facial features from the available data. The outer parts of the face were sometimes left in the image, making it possible to trace the scan back to an individual person. One explanation for this poor result is that the scans did not contain the expected contrast. The Freesurfer method expected true T_1 -weighted images and while the images available in the AHEAD dataset did include T_1 -weighted images, it is most likely that the noisy background caused the Freesurfer software to perform poorly. As a result, parts of the brain were incorrectly removed and parts of the face were left in the image. The brain mask allowed for correction for brain removal, but the leftover facial areas were insufficient to ensure complete anonymisation.

In an attempt to combat this, Nighres ([Huntenburg et al., 2018](#)) was used to detect the background of the image and remove it. This way, the Freesurfer software might more easily determine which part of the image belong to the face and which to the background. However, Nighres proved to faced issues with determining the background, most likely due to the subject region being noisy.

Further, the Freesurfer method used a Gaussian filter to smooth the edges between the removed face and the leftover scan. When the Freesurfer defacing removed part of the brain, the brain mask was used to restore the brain area. However, the brain has a sharp interface due to the masking operation. When applying Gaussian smoothing, part of the head will be smoothed. While the affected area is small, it might nonetheless be significant when training a reconstruction model on this data. All these concerns combined raised the need for another defacing method.

Finally, sufficient results were achieved with a dilation based method. Dilation of the brain mask allowed for complete removal of the face without losing any part of the brain or skull. This method did remove parts of the neck, which caused the scan to differ more from a non-defaced scan than would be preferred. However, this appears to be a necessary trade-off for the goal of releasing the dataset. Dilation is more flexible than a ready-made defacing algorithm when a brain mask is available, as it is easy to adjust the kernel size or number of iterations to exclusively remove a certain part of the face but leave features such as the skull intact. Using the brain mask as a base for this dilation can cause an uneven removal of the face, as the brain is somewhat round, and dilating this structure can cause the nose to be included in a scan while most of the eyes are left out. Fine-tuning a kernel size and the number of iterations is therefore of the highest importance, as the right kernel can soften this concern. Lastly, when making use of the brain mask, the same kernel size and number of iterations was proven to suffice for all subjects.

While this defacing method was created specifically to release the AHEAD dataset, it is important to consider other applications. In that light, a downside of using the brain mask for a defacing algorithm is that this method cannot be used for volumes of which a brain mask is not available. However, robust methods of segmenting the brain have been known since the mid-90s (see, for example, [Atkins and Mackiewich \(1998\)](#)). Today, methods for segmenting the brain from an MRI scan are widely available, easy to use, and provide high-quality results. Thus, if a brain mask is not yet available, one can easily be segmented. While the absence of a brain mask might be an extra hurdle, it is one that can be solved with readily available methods.

6.1.2 Adding Gaussian background noise

An MRI scan always contains noise, including the background region. Defacing using dilation removed this background, irretrievably changing the properties of the scan. To remedy this, noise was generated and added to the background. Results showed that the noise level of the additional noise was often computed to be higher

than that of the original noise. When the goal is to create a defaced scan that resembles an original scan as much as possible, this is far from ideal. One explanation for this overestimation is that the ring in k-space that the noise is sampled from could still contain some signal. While decreasing the width of the ring seemed like a valid solution, it still overestimated the noise level. Changing the ring size further came with additional drawbacks: decreasing the ring size too much and it becomes too small to accurately grasp the spread of the noise, increasing it too much and the estimation might be partly based on even more signal. It was deemed that the ring could not be much smaller in fear of the former.

Experiments showed that the noise level did not change much with a smaller ring, indicating that no real signal was present. A second explanation for the overestimation might come from the fact that the original scan was reconstructed using GRAPPA. GRAPPA reduces aliasing in parallel imaging accelerated MRI, but in the process will also alter the spatial distribution of noise (Aja-Fernández et al., 2014). This disturbance of noise can be estimated, usually by running Monte Carlo simulations with simulated Gaussian noise. In Parallel Imaging, the noise is influenced by the acceleration factor, but also by a coil-dependent factor known as the geometry (g)-factor (Breuer et al., 2009). This g-factor causes spatial variance and enhancement in the noise and is heavily dependent on the acceleration factor. An explanation for the overestimation in the noise of the defaced images lies in this g-factor, as the slice of k-space may have a different noise distribution than the box from which comparisons were made. Using a different slice of k-space would not be possible, as the likelihood of that slice containing signal would be higher. A solution would be to use data without motion correction and GRAPPA interpolation, as this would present the originally acquired noise distribution.

If, like in this work, this data is not available, a future alteration to the noise estimation could take the g-factor into account and correct for it. This way, a more even distribution of noise would be ensured. The presence of this g-factor also has the consequence that the added noise will most likely be more evenly distributed than the noise in the original scan. This is due to the fact that the added noise was not modified by the g-factor variances, as the standard deviation was taken from only one slice. While this can be seen as an improvement, as a more even distribution of noise is desired, it can also be a hindrance. With the intention of creating a scan that is similar to a non-defaced scan, this more evenly distributed noise is problematic as it differs from the properties of the original scan.

Defacing using dilation of the brain mask is highly versatile: it's easy to change the kernel size or run it for an extra iteration. On the other hand, the need to add background noise is a clear downside, as it is difficult to generate noise that is similar to the original scan. Unless precise noise measurements are made at the time of scanning, this will always be an estimation and might alter the properties of the image. That can be a concern when training a network to reconstruct the image in image space, as the network will learn an overestimation of noise in the background. As a consequence, the network might assume a higher noise level in the brain region, affecting the SNR and therefore the quality of the scan. Trying to correct the noise for overestimation was out of the scope of this work, but should be considered for future research. As a stand-in solution, both the scan with the added noise and the scan without added noise (i.e. a zero-filled background) will be released at the time of publishing of the dataset, such that research groups can themselves choose which scan to use when training a network.

A final and conclusive test to see whether the defaced images resemble a true MRI scan would be to train a network on the defaced data and compare testing results to the same network trained on the original images. This lies within the capabilities of the RIM (Lønning et al., 2019; Zhang et al., 2021), but training this network is time and resource consuming and thus was deemed out of the scope of this work.

Currently, the defacing method is performed on complex coil images. Due to the nature of this data, the defacing takes about 45 minutes per scan. Dilation itself only takes a few seconds and the calculation of noise are simple matrix operations. Most time is spent on loading and saving new versions of the data, as the dilated mask, the complex coil image with added noise, and the complex coil image without added noise are saved. However, since defacing only needs to be performed once, this is not a major cause for concern as of yet.

6.1.3 Dataset publication

At the time of writing, communications are still ongoing with Dataverse NL to ensure the release of the dataset. The prognosis, however, is that this will take place over the summer of 2021. A test environment has been set up and initial trial uploads were successful. However, the environment is not yet ready for release. The data will be released under a Creative Commons Attribution 4.0 International Public License. But as this licence is not part of Dataverse NL's default environment, manual changes are required prior to the release. The data currently also includes subject IDs which can lead back to individual persons. These numbers will be shuffled twice and the key will be purposely lost to ensure nothing can be traced back. Finally, decisions on which metadata to include are still pending. The aforementioned decisions and changes are important ones and need to be taken seriously.

6.2 Sharpness estimation along tissue boundaries

The results showed that sharpness estimation using two-fold k-means clustering is a valid predictor for determining the content-wise image quality of reconstructed R_2^* maps. Further, the sharpness estimation was able to determine an estimation of the sharpness for all relevant acceleration factors and identify differences between them. However, the method was unable to provide an estimation for some regions, e.g. the thalamus. This is most likely due to a lack of contrast with the surrounding tissue. This suspicion is supported by the performance of the estimation on the globus pallidus. This part of the brain is iron-rich and therefore shows high contrast on an R_2^* map. The sharpness estimation showed a clear upwards trend for this region, suggesting the estimation requires regions with a high enough contrast to adjacent regions to function properly. A lack of contrast will leave the method unable to make a sharpness estimate, which in turn causes some ROIs to be excluded from the analysis. The need for contrast is in line with expectations, as the method was originally created for T_1 weighted imaging, which has a higher contrast to noise ratio than an R_2^* map.

So far, the method can only be used to compare two scans at once: the ground truth and an accelerated scan. To facilitate a comparison between acceleration factors, the ground truth estimations needed to be fixed. This can result in the selection of sub-optimal clusters for higher acceleration rates, leading in turn to a sub-optimal sharpness estimation. A future implementation of the method could allow for comparing all acceleration factors at once. This would create a more robust method and allow for more insights into the differences between acceleration factors. On top of that, the subclusters for sharpness estimation are currently optimised for the scan with acceleration factor 3. With a comparison of all acceleration factors, this optimisation could be done for all scans, potentially increasing performance.

6.2.1 Trend analysis

Results showed an overall upwards trend for most regions. For some ROIs, namely the ventricles, the sharpness estimate showed to decrease instead of increase for a higher acceleration factor. This might be because a large number of estimations were deemed invalid due to the lack of contrast. As expected, all regions show an overall trend for acceleration factor as the RIM reconstructs a whole-brain image. If this sharpness estimation would show significant differences in regions, this could imply that the sharpness method would not work for R_2^* maps, e.g. due to lack of contrast. Further, a significant difference in sharpness estimation could indicate that the (AI) reconstruction method shows a bias towards some (easier to reconstruct) regions.

A paired t-test showed a significant difference between the sharpness of images with an acceleration factor of 3 and those with an acceleration factor of 12. This showed that the method is able to pick up a decrease in sharpness associated with an increase in acceleration factor. As this is in line with expectations, this sharpness estimation was therefore demonstrated as a valid method of determining the content-wise image quality of reconstructed images. Further, the method allows to gain insight into the quality of the contents of the image. While the community is devoted to improving the reconstruction quality, this method has been proven to reliably measure improvements in a given reconstruction method.

Chapter 7

Conclusion

Research in accelerated MRI reconstruction using Deep Learning continues to gain traction, with more and more research groups focusing on this topic. In that light, it is important to make MRI data available that can be used for this. To do so, this work focused on creating an anonymous dataset that would allow research groups without access to an MRI scanner to delve into the topic. The AHEAD dataset was successfully defaced by the use of dilation and was thus anonymised in order to prepare it for release to the public. The dilation algorithm was able to correctly remove any recognisable facial features, and with the use of Gaussian smoothing and generation of background noise, the image was made to largely resemble a true MRI scan.

The added noise was found to be significantly higher than the noise in the original scan. This was most likely due to the spatial variance of noise in GRAPPA reconstructed images. This overestimation of noise could potentially have a detrimental effect on the results of an MRI reconstruction method. The temporary solution is to release the data both with and without added noise, leaving the choice up to the researcher which scan to use. In future research, this difference in noise is something that should be corrected for, for example by taking this spatial variance of noise into account. This work showed that an ideal defacing method for versatile MRI data is not available as of yet. Both the Freesurfer defacing software and the dilation algorithm explored had their advantages, but also their drawbacks: one didn't remove all relevant facial areas, the other required longer runtimes and contained a difference in background noise. However, the defacing using dilation was able to anonymise MRI data without compromising the raw data too much. In addition, the actual release of the dataset has yet to take place, as talks concerning the accompanying metadata and copyright licence. The prognosis is that the actual release will take place somewhere in the summer of 2021.

Another aim of this work, and another way to advance research into accelerated MRI reconstruction, was to create a new method to allow for the assessment of the content-wise quality of reconstructed MR images. To do so, a method originally created for the assessment of motion-corrected T_1 -weighted images was re-purposed and adapted to assess R_2^* maps. With the use of two-fold k-means clustering, several regions of interest were investigated and the sharpness over the region edge was estimated. Results showed that this method is able to distinguish image quality between different acceleration factors. The sharpness estimation showed a significant difference in sharpness for a low and high acceleration factor, proving that the method can detect a decrease in sharpness for higher acceleration factors. This confirms the hypothesis that this method can be used to assess the content quality of reconstructed MRI images and can be used to advance reconstruction models.

The sharpness estimation method required sufficient contrast between regions in order to make an estimate. When applied to R_2^* maps, the method has shown to often lack this contrast, making the estimation less reliable. While R_2^* maps do contain ample regions with contrast to surrounding areas, meaning the method can certainly be used there, this need for contrast is a clear limitation of the sharpness estimation. Therefore, it is advisable to address this limitation in any future developments of this method. Another limitation is the ability to compare only two acceleration factors at once. To optimise the method, a comparison across all acceleration factors can lead to better and fairer estimations.

In conclusion, this work has delivered a valuable dataset that can be used to further research towards accelerated MRI reconstruction. Furthermore, a structural, quantitative method of assessing content-wise reconstruction quality was presented and successfully applied to R_2^* maps.

Bibliography

- Ahmad, R., Ding, Y., and Simonetti, O. P. (2015). Edge sharpness assessment by parametric modeling: application to magnetic resonance imaging. *Concepts in Magnetic Resonance Part A*, 44(3):138–149.
- Aja-Fernández, S., Vegas-Sánchez-Ferrero, G., and Tristán-Vega, A. (2014). Noise estimation in parallel mri: Grappa and sense. *Magnetic resonance imaging*, 32(3):281–290.
- Alkemade, A., Mulder, M. J., Groot, J. M., Isaacs, B. R., van Berendonk, N., Lute, N., Isherwood, S. J., Bazin, P.-L., and Forstmann, B. U. (2020). The amsterdam ultra-high field adult lifespan database (ahead): A freely available multimodal 7 tesla submillimeter magnetic resonance imaging database. *NeuroImage*, page 117200.
- Atkins, M. S. and Mackiewich, B. T. (1998). Fully automatic segmentation of the brain in mri. *IEEE transactions on medical imaging*, 17(1):98–107.
- Bazin, P.-L., Nijssse, H. E., van der Zwaag, W., Gallichan, D., Alkemade, A., Vos, F. M., Forstmann, B. U., and Caan, M. W. (2020). Sharpness in motion corrected quantitative imaging at 7t. *NeuroImage*, 222:117227.
- Bernstein, M. A., King, K. F., and Zhou, X. J. (2004). *Handbook of MRI pulse sequences*. Elsevier.
- Biasioli, L., Lindsay, A. C., Choudhury, R. P., and Robson, M. D. (2011). Loss of fine structure and edge sharpness in fast-spin-echo carotid wall imaging: measurements and comparison with multiple-spin-echo in normal and atherosclerotic subjects. *Journal of Magnetic Resonance Imaging*, 33(5):1136–1143.
- Bischoff-Grethe, A., Ozyurt, I. B., Busa, E., Quinn, B. T., Fennema-Notestine, C., Clark, C. P., Morris, S., Bondi, M. W., Jernigan, T. L., Dale, A. M., et al. (2007). A technique for the deidentification of structural brain mr images. *Human brain mapping*, 28(9):892–903.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Blaimer, M., Breuer, F., Mueller, M., Heidemann, R. M., Griswold, M. A., and Jakob, P. M. (2004). Smash, sense, pils, grappa: how to choose the optimal method. *Topics in Magnetic Resonance Imaging*, 15(4):223–236.
- Breuer, F. A., Kannengiesser, S. A., Blaimer, M., Seiberlich, N., Jakob, P. M., and Griswold, M. A. (2009). General formulation for quantitative g-factor calculation in grappa reconstructions. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 62(3):739–746.
- Caan, M. W., Bazin, P.-L., Marques, J. P., de Hollander, G., Dumoulin, S. O., and van der Zwaag, W. (2019). Mp2rageme: T1, t2*, and qsm mapping in one sequence at 7 tesla. *Human brain mapping*, 40(6):1786–1798.
- Caviedes, J. and Gurbuz, S. (2002). No-reference sharpness metric based on local edge kurtosis. In *Proceedings. International Conference on Image Processing*, volume 3, pages III–III. IEEE.
- Chen, Y., Yu, W., and Pock, T. (2015). On learning optimized reaction diffusion processes for effective image restoration. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5261–5269.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Damadian, R., Minkoff, L., Goldsmith, M., Stanford, M., and Koutcher, J. (1976). Field focusing nuclear magnetic resonance (fonar): visualization of a tumor in a live animal. *Science*, 194(4272):1430–1432.
- Griswold, M. A., Jakob, P. M., Heidemann, R. M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., and Haase, A. (2002). Generalized autocalibrating partially parallel acquisitions (grappa). *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 47(6):1202–1210.

- Hammernik, K., Klatzer, T., Kobler, E., Recht, M. P., Sodickson, D. K., Pock, T., and Knoll, F. (2018). Learning a variational network for reconstruction of accelerated mri data. *Magnetic resonance in medicine*, 79(6):3055–3071.
- Holmes, A. J., Hollinshead, M. O., O’keefe, T. M., Petrov, V. I., Fariello, G. R., Wald, L. L., Fischl, B., Rosen, B. R., Mair, R. W., Roffman, J. L., et al. (2015). Brain genomics superstruct project initial data release with structural, functional, and behavioral measures. *Scientific data*, 2(1):1–16.
- Hoult, D. I. and Bhakar, B. (1997). Nmr signal reception: Virtual photons and coherent spontaneous emission. *Concepts in Magnetic Resonance: An Educational Journal*, 9(5):277–297.
- Huntenburg, J. M., Steele, C. J., and Bazin, P.-L. (2018). Nighres: processing tools for high-resolution neuroimaging. *GigaScience*, 7(7):giy082.
- Hutchinson, M. and Raff, U. (1988). Fast mri data acquisition using multiple detectors. *Magnetic resonance in Medicine*, 6(1):87–91.
- Hyun, C. M., Kim, H. P., Lee, S. M., Lee, S., and Seo, J. K. (2018). Deep learning for undersampled mri reconstruction. *Physics in Medicine & Biology*, 63(13):135007.
- Jackson, J. A. and Langham, W. H. (1968). Whole-body nmr spectrometer. *Review of Scientific Instruments*, 39(4):510–513.
- Jain, A. K. (1989). *Fundamentals of digital image processing*. Prentice-Hall.
- Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2):825–841.
- Jenkinson, M. and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical image analysis*, 5(2):143–156.
- Jin, K. H., McCann, M. T., Froustey, E., and Unser, M. (2016). Deep convolutional neural network for inverse problems in imaging. *CoRR*, abs/1611.03679.
- Kwiat, D., Einav, S., and Navon, G. (1991). A decoupled coil detector array for fast image acquisition in magnetic resonance imaging. *Medical Physics*, 18(2):251–265.
- Lebel, C., Gee, M., Camicioli, R., Wieler, M., Martin, W., and Beaulieu, C. (2012). Diffusion tensor imaging of white matter tract evolution over the lifespan. *Neuroimage*, 60(1):340–352.
- Lønning, K., Putzky, P., Sonke, J.-J., Reneman, L., Caan, M. W., and Welling, M. (2019). Recurrent inference machines for reconstructing heterogeneous mri data. *Medical image analysis*, 53:64–78.
- Lustig, M., Donoho, D., and Pauly, J. M. (2007). Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 58(6):1182–1195.
- Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.-F., and Gruetter, R. (2010). Mp2rage, a self bias-field corrected sequence for improved segmentation and t1-mapping at high field. *Neuroimage*, 49(2):1271–1281.
- Mezrich, R. (1995). A perspective on k-space. *Radiology*, 195(2):297–315.
- Pruessmann, K. P., Weiger, M., Scheidegger, M. B., and Boesiger, P. (1999). Sense: sensitivity encoding for fast mri. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 42(5):952–962.
- Putzky, P. and Welling, M. (2017). Recurrent inference machines for solving inverse problems. *arXiv preprint arXiv:1706.04008*.
- Regulation, G. D. P. (2016). Regulation eu 2016/679 of the european parliament and of the council of 27 april 2016. *Official Journal of the European Union*. Available at: http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf (accessed 20 September 2017).
- Rieger, B. and van Veen, G. A. (2008). Method to determine image sharpness and resolution in scanning electron microscopy images. In *EMC 2008 14th European Microscopy Congress 1–5 September 2008, Aachen, Germany*, pages 613–614. Springer.

- Rinck, P. A. (2019). *Magnetic resonance in medicine: a critical introduction*. BoD–Books on Demand.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.
- Schlemper, J., Caballero, J., Hajnal, J. V., Price, A. N., and Rueckert, D. (2017). A deep cascade of convolutional neural networks for dynamic mr image reconstruction. *IEEE transactions on Medical Imaging*, 37(2):491–503.
- Schoormans, J., Strijkers, G. J., Hansen, A. C., Nederveen, A. J., and Coolen, B. F. (2020). Compressed sensing mri with variable density averaging (cs-vda) outperforms full sampling at low snr. *Physics in Medicine & Biology*, 65(4):045004.
- Siegelmann, H. T. and Sontag, E. D. (1991). Turing computability with neural nets. *Applied Mathematics Letters*, 4(6):77–80.
- Sodickson, D. K. and Manning, W. J. (1997). Simultaneous acquisition of spatial harmonics (smash): fast imaging with radiofrequency coil arrays. *Magnetic resonance in medicine*, 38(4):591–603.
- Szeliski, R. (2010). *Computer vision: algorithms and applications*. Springer Science & Business Media.
- Tona, K.-D., van Osch, M. J., Nieuwenhuis, S., and Keukens, M. C. (2019). Quantifying the contrast of the human locus coeruleus in vivo at 7 tesla mri. *PloS one*, 14(2):e0209842.
- Westbrook, C., Kaut Roth, C., and Talbot, J. (2011). *MRI in Practice*. John Wiley & Sons.
- Wilson, J. N. and Ritter, G. X. (2000). *Handbook of computer vision algorithms in image algebra*. CRC press.
- Yang, Y., Sun, J., Li, H., and Xu, Z. (2017). Admm-net: A deep learning approach for compressive sensing mri. *arXiv preprint arXiv:1705.06869*.
- Zbontar, J., Knoll, F., Sriram, A., Muckley, M. J., Bruno, M., Defazio, A., Parente, M., Geras, K. J., Katsnelson, J., Chandarana, H., et al. (2018). fastmri: An open dataset and benchmarks for accelerated mri. *arXiv preprint arXiv:1811.08839*.
- Zhang, C., Poot, D., Coolen, B., Vrenken, H., Bazin, P.-L., Forstmann, B., and Caan, M. (2021). A unified model for simultaneous reconstruction and r2* mapping of accelerated 7t data using the recurrent inference machine. *ISMRM*.
- Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R., and Rosen, M. S. (2018). Image reconstruction by domain-transform manifold learning. *Nature*, 555(7697):487–492.

Appendix A

Methods

AHEAD dataset

Sharpness levelmaps

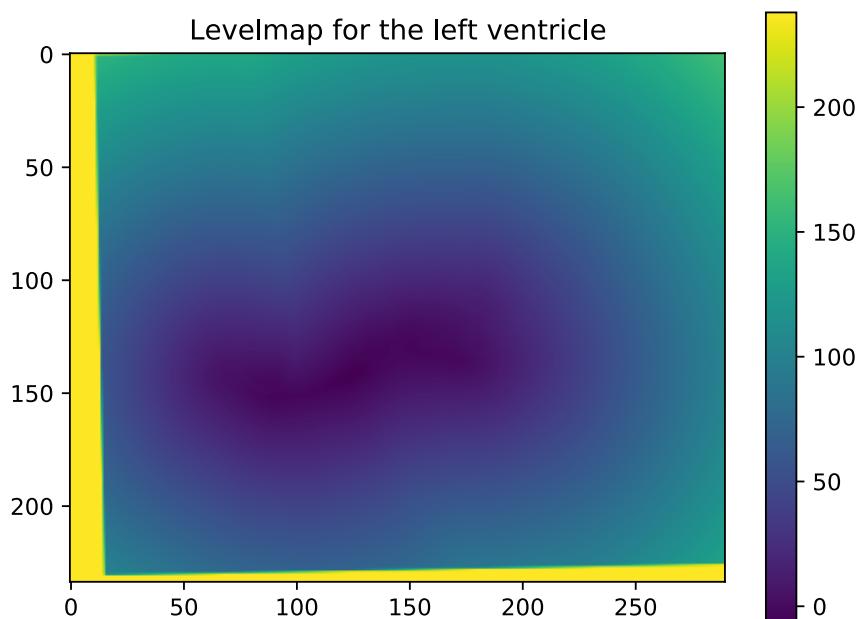


Figure A.1: Example of a coregistered levelmap used in the sharpness estimation. Values below zero indicate the voxel lies inside of the ROI while values above zero indicate the voxel lies outside of the ROI.

Age group (years)	Female	Male	Total
18–30	27	15	42
31–40	6	6	12
41–50	6	7	13
51–60	7	5	12
61–70	7	6	13
71–80	7	6	13
Total	60	45	105

Table A.1: Distribution of ages of the participants from the AHEAD dataset. Numbers taken from [Alkemade et al. \(2020\)](#).

Appendix B

Results

Sharpness profiles

Sharpness histogram

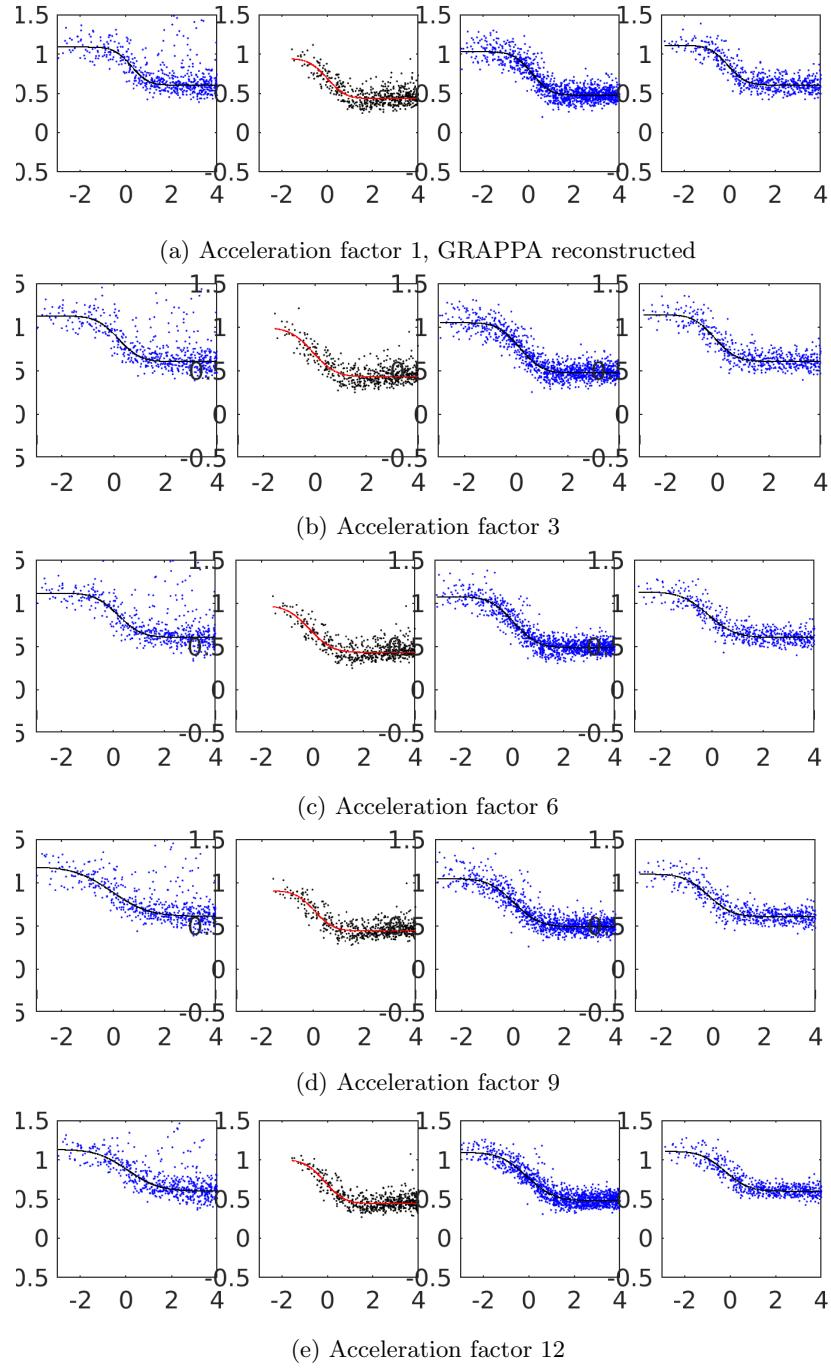


Figure B.1: Profiles from all acceleration factors of the sharpness estimation for the globus pallidus, showing whether a boundary is valid and contrasting enough to be included in the sharpness estimation.

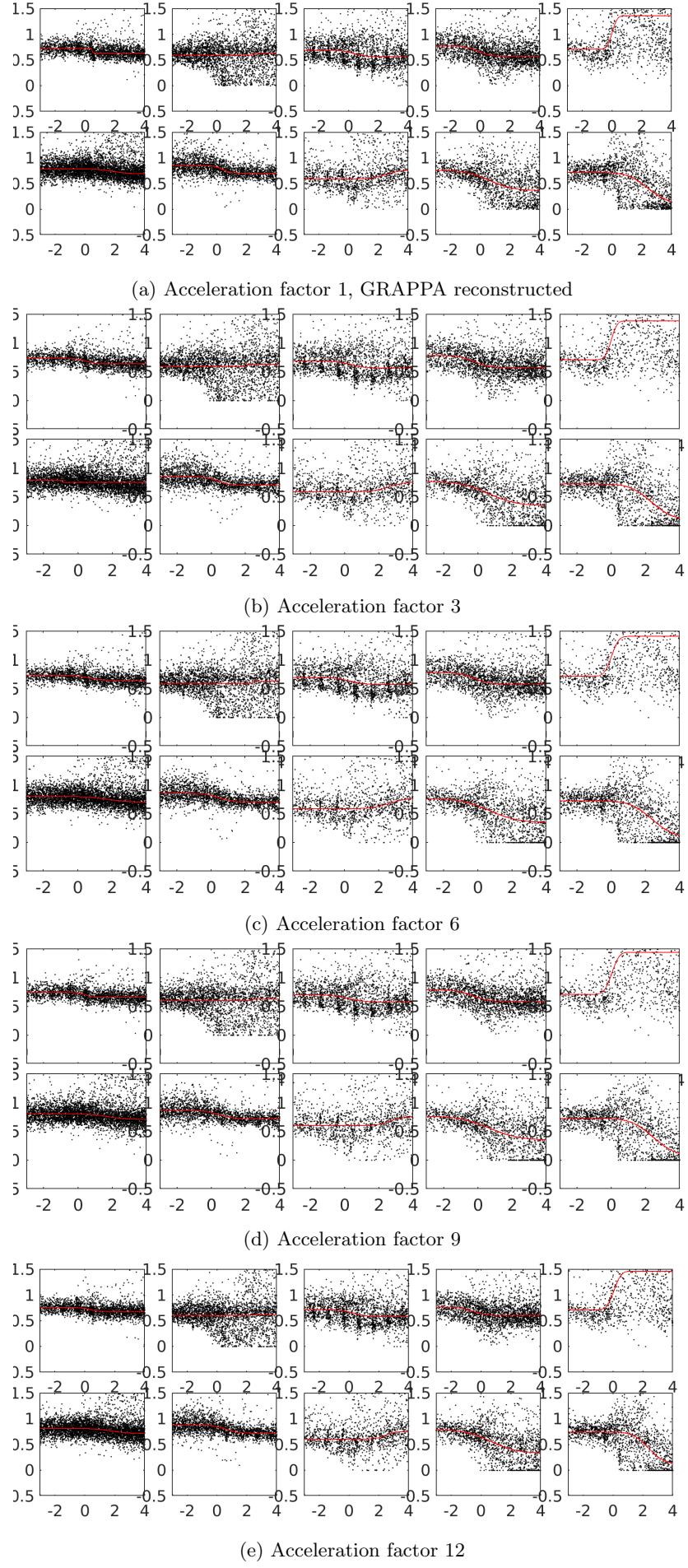
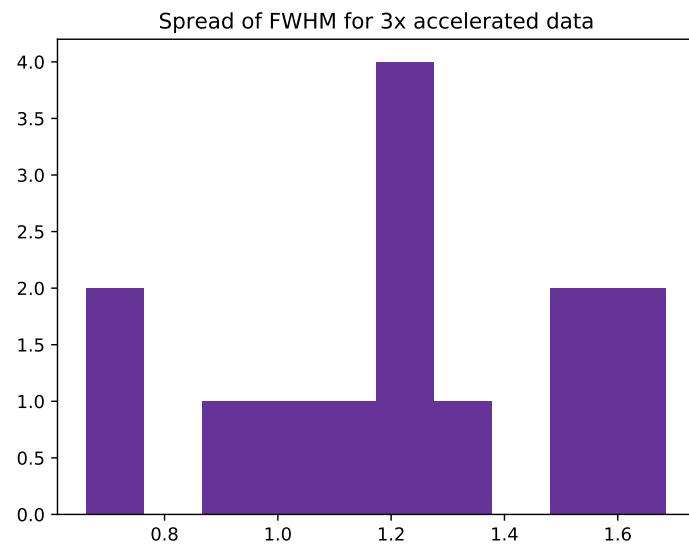
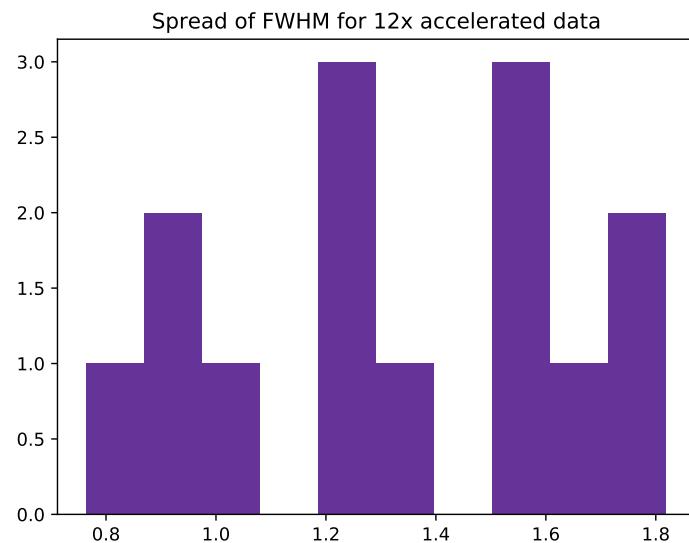


Figure B.2: Profiles from all acceleration factors of the sharpness estimation for the thalamus, showing whether a boundary is valid and contrasting enough to be included in the sharpness estimation.



(a) Distribution of FWHM for 3x accelerated data



(b) Distribution of FWHM for 12x accelerated data

Figure B.3: Histograms for initial data inspection showing the distribution of FWHM data for two different acceleration factors.