

CONVOLUTIONAL NEURAL NETWORKS FOR HEAD AND
NECK TUMOR SEGMENTATION IN MRI:
PROTOCOL AND SEQUENCE OPTIMIZATION

LARS BIELAK



INAUGURAL DISSERTATION
FOR THE DOCTORAL DEGREE
AT THE FACULTY OF MATHEMATICS AND PHYSICS,
ALBERT-LUDWIGS-UNIVERSITY FREIBURG

Lars Bielak: *Convolutional Neural Networks for Head and Neck Tumor Segmentation
in MRI: Protocol and Sequence Optimization*

Freiburg, 11 April 2022

SUPERVISORS:

Prof. Dr. Michael Bock

Prof. Dr. Dr. h.c. Jürgen Hennig

WRITTEN EXAMINERS:

Prof. Dr. Dr. h.c. Jürgen Hennig

Prof. Dr. Günter Reiter

ORAL EXAMINERS:

Prof. Dr. Dr. h.c. Jürgen Hennig

Prof. Dr. Jens Timmer

Prof. Dr. Oskar von der Lüche

DEAN OF THE FACULTY OF MATHEMATICS AND PHYSICS:

Prof. Dr. Michael Thoss

Because it's there.

— George Mallory

ABSTRACT

Radiation therapy of head and neck tumors requires precise tumor segmentations for an optimal treatment outcome. For treatment planning, tumors are best segmented on magnetic resonance imaging (MRI) data.

Today, convolutional neural networks (CNNs) can automate the tedious segmentation process for some tumors, with an excellent performance: However, this has not been realized in head and neck cancer so far. This thesis focuses on the connection of automatic tumor segmentation using CNNs to the MR image acquisition and protocol optimization. Therefore, the patient setup for head and neck cancer patients is improved to reduce artifacts in diffusion weighted MRI, and the effect of additional distortion correction in ADC parameter maps on the CNN performance is evaluated. For the segmentation, a feed-forward CNN with 2 pathways for high- and low-resolution features is realized for multi-parametric MRI data of head and neck cancer patients, achieving a segmentation Dice coefficient of up to 65%. The relative influence of each of 7 MRI input channels is quantified: It is shown that each of the 7 channels improves segmentation performance of the CNN, and that T_2^* has the largest overall influence, with a difference in segmentation performance to the reference network of 6%.

Furthermore, a new sequence is developed to simultaneously measure T_2 and ADC parameter maps. The multiecho spin echo sequence with interleaved diffusion blocks is designed to eliminate geometric distortion artifacts present in conventional diffusion weighted MRI using a highly undersampled radial readout in combination with a regularized iterative reconstruction. The sequence is extensively tested in simulations as well as in MRI phantoms, and first *in vivo* results are presented.

With the optimized MRI protocol and sequences for the segmentation performance of the CNN, the imaging time can be reduced and/or image quality can be improved.

ZUSAMMENFASSUNG

Bei der Bestrahlung von Kopf-Hals-Tumoren werden hochpräzise Tumorsegmentierungen benötigt, die auf Basis von magnetresonanztomographischen

(MRT)-Bildern erstellt werden, um bestmögliche Behandlungsergebnisse für die Patienten zu erzielen.

Neuste Entwicklungen zeigen, dass durch den Einsatz von neuronale Faltungsnetzwerken (engl.: convolutional neural networks, CNNs) der zeitaufwändige Segmentierungsprozess bei einigen Tumorentitäten automatisiert werden kann. Derartige Ergebnisse, die für einen klinischen Einsatz ausreichen würden, konnten bei Kopf-Hals-Tumor allerdings noch nicht gezeigt werden. Deshalb soll in dieser Arbeit das Zusammenspiel von automatischer Tumorsegmentierung mit Hilfe von CNNs, sowie der Bildaufnahme und Protokolloptimierung untersucht werden. Dazu wird zunächst der Patienten-Messaufbau optimiert um Bildartefakte in der Diffusionsbildgebung zu reduzieren. Weiterhin wird der Effekt der zusätzlichen Korrektur von Verzerrungsartefakten in der Diffusionsbildgebung auf die Qualität der CNN-Segmentierungen ausgewertet. Hierfür wird ein Feed-Forward CNN verwendet, welches mit Hilfe zwei separater Verarbeitungswege hoch- und niedrigaufgelöste Kontextinformationen von multiparametrischen MRT-Bildern verarbeitet. Mit dem CNN werden für die Kopf-Hals-Tumorpatienten Segmentierungsergebnisse (Dice-Sørensen-Coefficient, DSC) von bis zu 65% erreicht. Zusätzlich wird der relative Informationsgehalt von 7 MRT Eingabekanälen auf die Segmentierungsqualität quantifiziert: Für jeden der 7 Kanäle zeigt sich eine Verbesserung des Ergebnisses, wobei T_2^* mit einem Unterschied von 6% zum Referenznetzwerk den größten Einfluss aufweist.

Außerdem wird im Rahmen dieser Arbeit eine MRT-Sequenz zur gleichzeitigen Messung von T_2 - und ADC-Parameterkarten entwickelt. Diese basiert auf einer Multiecho Spin-Echo Sequenz mit zwischengeschalteten Diffusionsblöcken, um die starken Verzerrungsartefakte von konventioneller Diffusionsbildgebung im Kopf-Hals-Bereich zu umgehen. Zur Reduzierung der Aufnahmedauer wird eine stark unterabgetastete radiale Trajektorie verwendet, und die Bilder werden mit Hilfe eines regularisierten iterativen Algorithmus rekonstruiert. Die Sequenz wird in Simulationen sowie in MRT-Phantomen getestet und erste *in vivo* Ergebnisse werden gezeigt.

Mit Hilfe der daraus für die automatische Tumorsegmentierung abzuleitenden MRT-Protokoll- und Sequenzoptimierungen kann die Patienten-Messzeit verkürzt und/oder die Bildqualität signifikant verbessert werden.

PUBLICATIONS

FIRST-AUTHOR JOURNAL PAPER

- Bielak, L., Nicolay, N. H., Ludwig, U., Lottner, T., Rühle, A., Grosu, A.-L., and Bock, M.** Improvement of diffusion weighted MRI by practical B_0 homogenization for head & neck cancer patients undergoing radiation therapy. In: *Physica Medica* 97.1 (May 2022), pp. 59–65.
- Bielak, L., Wiedenmann, N., Berlin, A., Nicolay, N. H., Gunashekar, D. D., Hägele, L., Lottner, T., Grosu, A.-L., and Bock, M.** Convolutional Neural Networks for Head and Neck Tumor Segmentation on 7-Channel Multiparametric MRI: A Leave-One-out Analysis. In: *Radiation Oncology* 15.1 (July 2020), p. 181.
- Bielak, L., Wiedenmann, N., Nicolay, N. H., Lottner, T., Fischer, J., Bunea, H., Grosu, A.-L., and Bock, M.** Automatic Tumor Segmentation With a Convolutional Neural Network in Multiparametric MRI: Influence of Distortion Correction. en-gb. In: *Tomography: A Journal for Imaging Research* 5.3 (Sept. 2019), pp. 292–299.
- Bielak, L. and Bock, M.** Optimization of Diffusion Imaging for Multiple Target Regions Using Maximum Likelihood Estimation. en. In: *Current Directions in Biomedical Engineering* 3.2 (Sept. 2017), pp. 203–206.

CO-AUTHOR JOURNAL PAPER

- Gunashekar, D. D., **Bielak, L.**, Hägele, L., Oerther, B., Benndorf, M., Grosu, A.-L., Zamboglou, C., and Bock, M. Explainable AI for CNN-based prostate tumor segmentation in multi-parametric MRI correlated to whole mount histopathology. In: *Radiation Oncology* 17 (Apr. 2022), p. 65.
- Oerther, B., Buren, M. V., Klein, C. M., Kirste, S., Nicolay, N. H., Sprave, T., Spohn, S., Gunashekar, D. D., Hägele, L., **Bielak, L.**, Bock, M., Grosu, A.-L., Bamberg, F., Benndorf, M., and Zamboglou, C. Predicting Biochemical Failure in Irradiated Patients With Prostate Cancer by Tumour Volume Measured by Multiparametric MRI. en. In: *In Vivo* 34.6 (Jan. 2020), pp. 3473–3481.

- Wiedenmann, N., Grosu, A.-L., Büchert, M., Rischke, H. C., Ruf, J., **Bielak, L.**, Majerus, L., Rühle, A., Bamberg, F., Baltas, D., Hennig, J., Mix, M., Bock, M., and Nicolay, N. H. **The Utility of Multiparametric MRI to Characterize Hypoxic Tumor Subvolumes in Comparison to FMISO PET/CT. Consequences for Diagnosis and Chemoradiation Treatment Planning in Head and Neck Cancer.** en. In: *Radiotherapy and Oncology* 150 (Sept. 2020), pp. 128–135.
- Wiedenmann, N., Bunea, H., Rischke, H. C., Bunea, A., Majerus, L., **Bielak, L.**, Protopopov, A., Ludwig, U., Büchert, M., Stoykow, C., Nicolay, N. H., Weber, W. A., Mix, M., Meyer, P. T., Hennig, J., Bock, M., and Grosu, A. L. **Effect of Radiochemotherapy on T2* MRI in HNSCC and Its Relation to FMISO PET Derived Hypoxia and FDG PET.** In: *Radiat Oncol* 13 (Aug. 2018).

FIRST-AUTHOR CONFERENCE ABSTRACTS

- Bielak, L.**, Lottner, T., and Bock, M. **Radial Fast Spin Echo MRI with Compressed Sensing for Simultaneous ADC and T2 Mapping.** In: *Proc. Int. Soc. Mag. Reson. Med.* ISMRM 29th Annual Meeting & Exhibition. Vol. 29. Virtual Meeting, May 20, 2021, p. 4166.
- Bielak, L.**, Wiedenmann, N., Berlin, A., Hägele, L., Lottner, T., Gross, S., Grosu, A.-L., and Bock, M. **Evaluating Noise Robustness of CNN-based Head&Neck Tumor Segmentations on Multiparametric MRI Data.** In: *Proc. Int. Soc. Mag. Reson. Med.* ISMRM 28th Annual Meeting & Exhibition. Vol. 28. Virtual Meeting, Aug. 10, 2020, p. 0269.
- Bielak, L.**, Lottner, T., Wiedenmann, N., Bunea, H., Grosu, A.-L., and Bock, M. **Readout-segmented vs. Single Shot Diffusion MRI for Radiation Therapy Planning in Head and Neck Tumors.** In: *Proc. Int. Soc. Mag. Reson. Med.* ISMRM 27th Annual Meeting & Exhibition. Vol. 27. Montréal, QC, Canada, Apr. 26, 2019, p. 2359.
- Bielak, L.**, Wiedenmann, N., Lottner, T., Bunea, H., Grosu, A.-L., and Bock, M. **Quantifying Information Content of Multiparametric MRI Data for Automatic Tumor Segmentation using CNNs.** In: *Proc. Int. Soc. Mag. Reson. Med.* ISMRM 27th Annual Meeting & Exhibition. Vol. 27. Montréal, QC, Canada, Apr. 26, 2019, p. 2339.
- Bielak, L.**, Bunea, H., Wiedenmann, N., Grosu, A.-L., and Bock, M. **Image-based Multi-Scale Distortion Correction: Application to Diffusion Imaging.** In: *Proc. Int. Soc. Mag. Reson. Med.* ISMRM 26th Annual Meeting & Exhibition. Vol. 26. Paris, France, June 1, 2018, p. 1633.

Bielak, L., Bunea, H., Wiedenmann, N., Grosu, A.-L., and Bock, M. **The Weakest Link in the Chain: How MR Data Quality influences Convolutional Neural Network Performance.** In: *Proc. Int. Soc. Mag. Reson. Med.* ISMRM 26th Annual Meeting & Exhibition. Vol. 26. Paris, France, June 1, 2018, p. 2720.

CO-AUTHOR CONFERENCE ABSTRACTS

Gunashekar, D. D., **Bielak, L.,** Berlin, A., Hägele, L., Oerther, B., Benndorf, M., Grosu, A.-L., Zamboglou, C., and Bock, M. **Explainable AI for CNN-Based Prostate Tumor Segmentation in Multi-Parametric MRI Correlated to Whole Mount Histopathology.** In: *Proc. Int. Soc. Mag. Reson. Med.* Vol. 29. Virtual Meeting, May 2021, p. 0819.

Lottner, T., Fischer, J., Reiss, S., **Bielak, L.,** Heidt, T., Özen, A. C., Thielmann, J., von zur Mühlen, C., and Bock, M. **Multi-Echo GRASP for Cardiac T2*-Relaxometry.** In: *Proc. Int. Soc. Mag. Reson. Med.* Vol. 29. Virtual Meeting, May 2021, p. 3602.

Lottner, T., Reiss, S., Haas, L. V., Erhardt, J., Özen, A. C., **Bielak, L.,** Schüttler, M., and Bock, M. **RF Heating of a 64-Channel Implantable Electrode Grid: Simulation and Temperature Measurements.** In: *Proc. Int. Soc. Mag. Reson. Med.* Vol. 28. Virtual Meeting, Aug. 2020, p. 4199.

Lottner, T., Erhardt, J. B., Reichert, A., Özen, A. C., **Bielak, L.,** Schuettler, M., and Bock, M. **Avoiding Systematic Errors of Fiber-Optic Temperature Probes: MR Thermometry for Safety Assessment of Small Planar Metallic Structures.** In: *Proc. Int. Soc. Mag. Reson. Med.* Vol. 27. Montréal, QC, Canada, Apr. 2019, p. 4185.

Wiedenmann, N., Bunea, H., Rischke, H., Bunea, A., Nicolay, N. H., Majerus, L., **Bielak, L.,** Protopopov, A., Ludwig, U., Büchert, M., Stoykow, C., Mix, M., Bock, M., and Grosu, A. **EP-2030 Multiparametric MRI and FMISO PET in HNSCC and Its Relation with Outcome.** In: *Radiotherapy and Oncology.* ESTRO 38, 26-30 April 2019, Milan, Italy 133 (Apr. 2019), S1114–S1115.

CONTENTS

Introduction	1
I THEORY	
1 NUCLEAR MAGNETIC RESONANCE	7
1.1 Nuclear Spin	7
1.2 Spin Ensembles	8
1.3 Excitation	10
1.4 Relaxation	11
1.5 Resonance Effects - The Echo	12
2 IMAGING	19
2.1 k-Space - Measuring in the Fourier Domain	19
2.1.1 Image Space and k-Space	22
2.2 Pulse Sequences	23
2.2.1 Carr-Purcell-Meiboom-Gill Sequence	23
2.2.2 Spin Echo EPI	25
2.3 Accelerated Imaging	26
2.3.1 Parallel Imaging	26
2.3.2 Undersampling	27
3 QUANTITATIVE MRI	29
3.1 T_2 Measurements	29
3.2 Diffusion	31
3.2.1 Molecular Self-Diffusion	31
3.2.2 Diffusion Measurements	32
3.2.3 Artifacts in DW-EPI	34
4 ARTIFICIAL INTELLIGENCE	39
4.1 The Perceptron	40
4.2 Artificial Neural Networks	41
4.3 Convolutional Neural Networks	43
4.4 Network Architecture	45
II MATERIALS AND METHODS	
5 AUTOMATED TUMOR SEGMENTATION	51
5.1 FMISO Study	51
5.1.1 MRI Data of Head and Neck Tumors	51
5.1.2 Ground Truth Segmentations	52

5.1.3	Measurement Setup	55	
5.1.4	Measurement Protocol	57	
5.2	Single Shot EPI vs. Readout Segmented EPI		58
5.3	B_0 Field Homogenization	59	
5.4	Segmentation Algorithm	61	
5.4.1	CNN Architecture	61	
5.4.2	Data Pre-Processing	63	
5.5	Distortion Correction	64	
5.5.1	Retrospective Distortion Correction		64
5.5.2	Available Data	66	
5.5.3	Data Pre-Processing	66	
5.5.4	CNN Architecture	67	
5.5.5	Experiment Design	68	
5.5.6	Computational Resources	69	
5.6	Channel-Wise Importance Analysis		69
5.6.1	Available Data	69	
5.6.2	Data Pre-Processing	70	
5.6.3	CNN Architecture	71	
5.6.4	Experiment Design	74	
5.6.5	Computational Resources	75	
6	SIMULTANEOUS ADC AND T_2 MAPPING		77
6.1	Sequence Design	77	
6.2	Image and Parameter Map Reconstruction		80
6.2.1	Non-Uniform Fast Fourier Transform		80
6.2.2	Model-Based Reconstruction	80	
6.2.3	Model-Interleaved Reconstruction		82
6.2.4	Density Compensation	83	
6.2.5	Echo Selection for Reconstruction		83
6.2.6	Coil Sensitivity Maps	84	
6.2.7	Gradient Delay Correction	84	
6.3	Simulation	84	
6.3.1	Bloch Simulation	85	
6.3.2	Analytical Phantom Reconstruction		86
6.4	Measurements	88	
III RESULTS			
7	AUTOMATED TUMOR SEGMENTATION		93
7.1	Single Shot EPI vs. Readout Segmented EPI		93
7.2	B_0 Field Homogenization	94	
7.3	Distortion Correction	97	
7.3.1	Retrospective Distortion Correction		97

7.3.2	Segmentation Results	98
7.4	Channel-Wise Importance Analysis	100
8	SIMULTANEOUS ADC AND T_2 MAPPING	105
8.1	Simulation	105
8.1.1	Bloch Simulation	105
8.1.2	Analytical Phantom Reconstruction	107
8.2	Measurements	108
IV DISCUSSION AND OUTLOOK		
9	DISCUSSION	119
9.1	Automated Tumor Segmentation	119
9.1.1	B_0 Field Homogenization	119
9.1.2	Distortion Correction	120
9.1.3	Channel-Wise Importance Analysis	123
9.1.4	Conclusion	125
9.2	Simultaneous ADC and T_2 Mapping	125
10	OUTLOOK	129
V APPENDICES		
11	SUPPLEMENTARY MATERIAL	133
11.1	Automated Tumor Segmentation	133
11.2	Simultaneous ADC and T_2 Mapping	145
	BIBLIOGRAPHY	149

ACRONYMS

AI	Artificial Intelligence
ANN	Artificial Neural Network
ADC	Apparent Diffusion Coefficient
BG	Background
BW	Bandwidth
CG	Conjugate Gradient
CNN	Convolutional Neural Network
CPMG	Carr-Purcell-Meiboom-Gill
CSF	Cerebrospinal Fluid
CT	Computed Tomography
DCE	Dynamic Contrast Enhanced
DSC	Dice Sørensen Coefficient
DW	Diffusion Weighted
DWI	Diffusion Weighted Imaging
EPI	Echo Planar Imaging
ETL	Echo Train Length
FFT	Fast Fourier Transform
FID	Free Induction Decay
FSE	Fast Spin Echo
FOV	Field Of View
GPU	Graphical Processing Unit
GT	Ground Truth

GTV	Gross Tumor Volume
GTV-T	Gross Tumor Volume – Tumor
GTV-Ln	Gross Tumor Volume – Lymph Node Metastases
HNSCC	Head and Neck Squamous Cell Carcinoma
LOO	Leave-One-Out
MLP	Multilayer Perceptron
MRI	Magnetic Resonance Imaging
NMR	Nuclear Magnetic Resonance
NN	Neural Network
NUFFT	Non-Uniform Fast Fourier Transform
PET	Positron Emission Tomography
ReLU	Rectified Linear Unit
RF	Radio-Frequency
ROI	Region of Interest
rsEPI	Readout-Segmented Echo Planar Imaging
RT	Radiation Therapy
SE	Spin Echo
SGD	Stochastic Gradient Descent
SNR	Signal to Noise Ratio
ssEPI	Single-Shot Echo Planar Imaging
TE	Echo Time
TR	Repetition Time
TSE	Turbo Spin Echo (same as FSE)
TV	Total Variation

INTRODUCTION

Since Paul Lauterbur [Lau73] and Peter Mansfield [Man77] introduced magnetic resonance imaging (MRI) in the 1970s, MRI has become one of the most important and versatile diagnostic imaging tools. MRI is a noninvasive imaging technique that is free of ionizing radiation, and it offers capabilities of generating superior soft tissue contrast with arbitrary tomographic slice positioning.

The free choice between several image contrasts is a characteristic feature of MRI, which allows the acquisition of various tissue and disease-specific information. Clinically, the most common contrasts are based on the T_1 and T_2 relaxation times, in combination with the proton density. In many cases, contrast agents such as gadolinium-chelates are used to shorten the relaxation times, which can increase the contrast, for example, in well-perfused tumors.

Moreover, MRI can be used to quantitatively assess the underlying physical parameters that contribute to the measured signal. For example, the measurement of the molecular self-diffusion, as described by Stejskal and Tanner for the nuclear magnetic resonance (NMR) experiment [ST65], is an essential part of many clinical imaging protocols. There, diffusion information is a surrogate parameter for changes in cell structure, such as cell density. Other parameters, such as T_1 , T_2 , T_2^* , or quantitative perfusion using dynamic contrast enhanced (DCE) MRI can also be acquired. However, due to the longer acquisition times, these quantitative methods are often not implemented in clinical routine.

With the large variability in image contrasts, standardized protocols are needed, such as BI-RADS [MCL13; MSK21] for breast and PI-RADS [Wei16; ST16] for prostate cancer imaging. With the potential to acquire various functional and anatomical information in a single exam, MRI has become one of modern medicine's central cornerstones in tumor imaging.

For cancer patients, treatment options include surgery, chemotherapy, or radiation therapy (RT), or a combination thereof. In RT, ionizing radiation (high-energy photons, electron beams or alpha-radiation) is delivered to the lesion to initiate the cell death in the tumor. In RT, photon beams with energies in the range of 4-20 MeV are created by a linear accelerator (LINAC). Modern RT systems can shape the photon beam and irradiate from different angles to achieve a high radiation dose within the target structure, while sparing the surrounding healthy tissue. This dose plan is calculated from attenuation coefficients of the tissue, which can be derived from CT images. To achieve the best treatment result, an accurate segmentation of a lesion (i. e., the tumor) is

important in RT, which can best be achieved on MR images due to their high soft-tissue contrast.

In this thesis, MR imaging and post processing methods are developed for head and neck tumors. The vast majority of tumors (> 90%) in the region are head and neck squamous cell carcinoma (HNSCC). HNSCC is a cancer of the mucosa and can appear in any of the anatomical regions of the pharynx (oro- and hypopharynx), oral cavity, and the larynx. As it can spread to the lymphatic system, imaging down to the shoulder level is necessary for a complete diagnosis. Major risk factors for the development of head and neck cancer are tobacco or alcohol abuse, increasing the risk by 5-25 fold [Wys13; Blo88]. Additionally, a chronic infection with viruses of the human papilloma family can increase the risk of HNSCC [Ang10].

Although HNSCC is well curable in early stages, e.g. with radiation therapy, about 50% of the patients already developed advanced tumors, resulting in an average survival rate of only 52-64% [GP06], and more than 380.000 deaths per year, which makes head and neck cancer the 9th most common cause for cancer related-death.

MRI protocols of the head and neck area include T_1 -weighted images that should have excellent contrast between fat and the lesion, which typically shows similar signal behavior as muscle and considerably lower intensity than fat [RSH08]. T_2 -weighted images with fat suppression are acquired for a better delineation of the lesion from the surrounding tissue. Additionally, T_1 -weighted contrast-enhanced images are acquired, as the increased vascularity of the tumor leads to an early signal enhancement with respect to the surrounding healthy tissue. Lymph node metastases are detected by their change in size and shape, and, if necrotic, can present with a reduced signal in T_1 -weighted and an increased signal in T_2 -weighted imaging [RSH08].

In RT planning, contours of tumor and lymph node metastases are created manually based on all available images. This manual segmentation task is very time-consuming, taking up to two hours per patient for a complete segmentation. Additionally, the manual tumor delineations are subject to a large inter-observer variability [HST10; Gue19; vGN19], leading to different treatment outcomes depending on the performing radiation oncologist.

With the success of computer vision starting in the early 2000s, artificial intelligence based image classification and segmentation have become feasible. Automatic segmentation algorithms are trained on large databases of segmented images, allowing them to extrapolate to unseen data after a training phase. Thus, their ability to generalize is critically dependent on the amount and variability of the data used during training. Therefore, the availability of large public databases, including a million images with 1000 different class labels,

marked the breakthrough of neural networks in the computer vision community [LBH15; KSH12].

A similar development is seen for medical images, despite significant differences between tomographic and natural images. Recently, neural networks have been introduced into clinical workflows as assistance systems with the potential to drastically reduce the time for segmentation. However, as opposed to the vast publicly available natural image databases, medical image databases first needed to be established and carefully curated. Such publicly available databases do now exist for certain cancer entities, e. g. prostate [Lit17a], breast [New21] or brain tumors [Bak17]. Still, they are limited to specific diseases and the imaging modalities and contrasts included. Therefore, automatic segmentation of less frequent tumors or studies including MRI contrasts that are not commonly used in clinical routine cannot rely on large image databases.

Most of the current research work aims to develop more efficient or more precise neural networks, disregarding the choice of image contrast in MRI. The main goal of this thesis is to close this gap between improvements in computer vision and MR method development. Hence, a method to find optimal MR image contrasts is described that maximizes the segmentation performance of state of the art convolutional neural networks (section 5.6 and 7.4). Furthermore, improvements in the acquisition and post-processing of diffusion weighted (DW) MR images used for automatic tumor segmentation are presented. The patient setup for head and neck MRI was improved to increase the quality as well as the quantity of DW-MRI, which were used to train segmentation algorithms (sections 5.3 and 7.2). Next, additional post-processing to correct for distortion artifacts in DW-MRI was explored as a means to improve the segmentation quality of a neural network (sections 5.5 and 7.3).

From these results, the need for faster, artifact-free acquisition of different parameters becomes apparent, and therefore a new MRI sequence was proposed (chapter 6). The sequence was designed for simultaneous quantitative T_2 and diffusion parameter mapping, which is insensitive to geometric distortion as seen in conventional DW-MRI. In chapter 8, simulations and first *in vivo* results of the proposed sequence are presented.

Part I

THEORY

NUCLEAR MAGNETIC RESONANCE

Magnetic Resonance Imaging (MRI) is based on the resonant excitation of nuclear magnetic spin states. In nearly all clinical applications, the relevant nucleus is the hydrogen atom. A rigorous physical description of the MRI signal behavior requires a thorough understanding of atomic physics on a quantum mechanical level - however, most effects observed by and necessary for MRI sequences can be described by classical models. This chapter introduces nuclear magnetic resonance (NMR), i. e. basics of excitation and relaxation of spin ensembles, which lies the groundwork for MRI. This chapter is based on textbooks from Levitt [Levo8] and Brown et. al. [Bro14], which are also excellent resources for further reading.

1.1 NUCLEAR SPIN

Spin is an inherently quantummechanical internal degree of freedom of elementary particles, similar to charge. In general, spin is always associated with a total spin quantum number I , analogous to the angular momentum quantum numbers L , with a total angular momentum

$$I_{\text{tot}} = \hbar \sqrt{I(I+1)} \quad . \quad (1.1)$$

Due to the quantum mechanical nature of spin, the spin quantum number I is associated with $2I + 1$ eigenstates m_i along their quantization axis, with

$$m_i \in [-I, -I + \frac{1}{2}, \dots, I - \frac{1}{2}, I] \quad . \quad (1.2)$$

A nucleon, i. e. a proton or neutron, consists of three permanent elementary particles (quarks) with spin $I = \frac{1}{2}$ and a "sea" of force particles (gluons) and additional quarks. Neutrons and protons combine two quarks with antiparallel spin and therefore have a net spin of $\frac{1}{2}$. Combining protons and neutrons can create atoms with larger spin, both integer (bosons) and half integer (fermions). Here, only hydrogen nucleus (^1H) is discussed, which is (almost) exclusively

used for imaging in the human body due to its vast natural abundance in tissue water.

Measurement of the spin angular momentum along any axis can only yield the eigenstates of the wave function along that axis. For ^1H , $I = \frac{1}{2}$, such that $m_i = \pm\frac{1}{2}$ (in units of \hbar).

Spin is coupled to the magnetic moment $\hat{\mu}$ of a particle by a linear relationship

$$\hat{\mu} = \gamma \hat{I} \quad , \quad (1.3)$$

where the hat symbol ($\hat{\cdot}$) marks the argument to be a quantum mechanical operator. The gyromagnetic ratio γ is dependent on the atom and can be positive or negative. For hydrogen, $\frac{\gamma}{2\pi} = 42.577 \text{ MHz/T}$. Spin and magnetization have a directionality, and a particle is referred to as "polarized" along the axis of magnetization.

To simplify notation and to be consistent with the literature resources, energies will be given in units of \hbar if not stated otherwise.

Given an external magnetic field B_0 in z direction, there are two eigenstates of the z -angular momentum operator \hat{I}_z , called $|\alpha\rangle$ or *up* and $|\beta\rangle$ or *down*. The external field leads to an energy difference of the two states, the Zeeman splitting, with energy levels given by

$$\begin{aligned} \hat{\mathcal{H}}_0|\alpha\rangle &= +\frac{1}{2}\omega_0|\alpha\rangle \\ \hat{\mathcal{H}}_0|\beta\rangle &= -\frac{1}{2}\omega_0|\beta\rangle \quad . \end{aligned} \quad (1.4)$$

Here, $\hat{\mathcal{H}}_0$ is the Hamilton operator (energy operator), and therefore $\pm\frac{1}{2}\omega_0$ are the energy eigenvalues of the two eigenstates. ω_0 is given by

$$\omega_0 = \gamma B_0 \quad (1.5)$$

and describes the precession frequency of the magnetic moment about the B_0 -axis and is called the *Larmor frequency*.

1.2 SPIN ENSEMBLES

In an NMR or MRI experiment individual spins are not considered, but rather the combined action of many independent ($\approx 10^{22}$) molecules is observed. Therefore, the spin dynamics are described by the spin density operator

$$\hat{\rho} = \begin{pmatrix} \rho_{\alpha\alpha} & \rho_{\alpha\beta} \\ \rho_{\beta\alpha} & \rho_{\beta\beta} \end{pmatrix} \quad , \quad (1.6)$$

where $\rho_{\alpha\alpha}$ and $\rho_{\beta\beta}$ denote the populations of the eigenstates $|\alpha\rangle$ and $|\beta\rangle$ and $\rho_{\alpha\beta}$ and $\rho_{\beta\alpha}$ denote coherence states. The density operator averages the states of all spins in the ensemble. Physically, the difference in populations is a measure for the magnetization of the spin ensemble along an external field axis, while the coherences can be interpreted as the magnetization state of the ensemble in the transversal plane, relative to an external field direction.

Thus, the thermal equilibrium magnetization can be derived from the spin density operator, given that in equilibrium the coherence states $\rho_{\alpha\beta} = \rho_{\beta\alpha} = 0$ and that the energy eigenstates are populated according to the Boltzmann statistics:

$$\begin{aligned}\rho_{\alpha\alpha} &= \frac{e^{-\frac{\hbar\omega_0}{2k_B T}}}{Z} \\ \rho_{\beta\beta} &= \frac{e^{+\frac{\hbar\omega_0}{2k_B T}}}{Z} \\ Z &= e^{-\frac{\hbar\omega_0}{2k_B T}} + e^{+\frac{\hbar\omega_0}{2k_B T}} .\end{aligned}\tag{1.7}$$

The difference of the two populations denotes the amount of particles, which effectively contribute to the net magnetization \vec{M}_0 , which is normalized to the volume V and the total number of particles N :

$$|\vec{M}_0| = \frac{N\gamma\hbar}{V} (\rho_{\alpha\alpha} - \rho_{\beta\beta}) .\tag{1.8}$$

At room temperature and typical clinical MRI field strengths of 1.5-3 T, the factor $(\rho_{\alpha\alpha} - \rho_{\beta\beta})$ is in the order of magnitude of 10^{-6} , which means that only about 1 ppm of the available spins can be detected in an MRI experiment.

It is possible to translate the spin density operator in a vector representation of the magnetization, with M_z linked to the populations, and M_x and M_y linked to the coherence states. This representation conveniently links to the classical description of a moment \vec{M} in an external field \vec{B} , which is described by the differential equation

$$\frac{d\vec{M}}{dt} = \gamma\vec{M} \times \vec{B} .\tag{1.9}$$

Equation (1.9) describes the precession of the magnetization vector \vec{M} about the axis defined by \vec{B} (here, the z-axis). The precession frequency is given by ω_0 , as shown in figure 1.1 a).

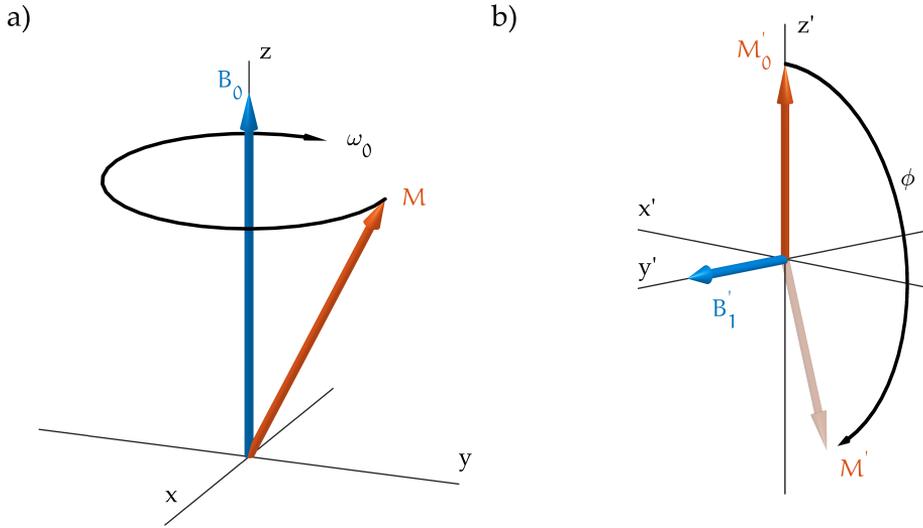


Figure 1.1: Dynamics of the magnetization vector in the presence of an external field. a) Precession of M around B_0 with the Larmor frequency ω_0 . b) In the rotating frame, M' is turned around the y' -axis by the transmit field B'_1 . In the laboratory frame, B'_1 rotates with ω_0 , such that it appears stationary in the rotating frame.

1.3 EXCITATION

In the equilibrium state, \vec{M} is parallel or antiparallel ($\gamma < 0$) to the external field \vec{B}_0 , and $\frac{d\vec{M}}{dt} = 0$. However, as suggested by equation (1.9), additional magnetic fields can be introduced to rotate the magnetization from the equilibrium state. This process is called radio frequency (RF) excitation and lies at the basis of any NMR experiment.

The dynamics of the magnetization vector can be simplified if the frame of reference is rotating with the same angular frequency as the precession frequency, called the *rotating frame*. A basis change into the rotating frame of reference with angular frequency ω_0 creates a virtual magnetic field, which is equal and opposite to B_0 , resulting in a net external magnetic field of $B'_0 = 0$. Generally,

$$\vec{B}' = \vec{B} + \frac{\vec{\Omega}}{\gamma} \quad , \quad (1.10)$$

where $\vec{\Omega}$ is the angular velocity of the rotating frame and \vec{B}' is the effective magnetic field in the rotating frame. From here, the rotating frame of reference is marked by the prime ($'$) sign. In the example above, $\vec{\Omega} = -\gamma B_0 \vec{e}_z$ and $\vec{B}_0 = B_0 \vec{e}_z$.

If an additional magnetic field \vec{B}_1 is applied with the same angular frequency ω_0 in the transverse x - y plane, it will appear stationary in the rotating frame, i. e. $\vec{B}'_1 = \text{const.}$ However, equation (1.9) still holds, and \vec{M}' will now precess about \vec{B}'_1 . Typically, \vec{B}_1 is applied perpendicular to the B_0 -field, and, without loss of generality, $\vec{B}'_1 = B'_1 \vec{e}'_y$. Consequently, \vec{M}' is rotated from its equilibrium position along the z' -axis around the y' -axis. The angle ϕ of rotation can be taken from the solution of the differential equation:

$$\begin{aligned} M'_z(t) &= M'_0 \text{Re}(e^{i\phi}) \\ M'_x(t) &= M'_0 \text{Im}(e^{i\phi}) \\ M'_y(t) &= 0 \quad \text{and} \\ \phi &= \gamma \int_{t_0}^{t_1} B'(t) dt \quad . \end{aligned} \tag{1.11}$$

The angle ϕ is commonly called the *flip angle*, and is governed by the amplitude and duration of the B_1 field. The effect of the \vec{B}'_1 field on the magnetization is depicted in figure 1.1 b).

A common simplification of the magnetization is performed by splitting the vector \vec{M} into two parts: the transversal component $M_{\perp} = \|(M_x, M_y)\|$ and the longitudinal component $M_{\parallel} = M_z$. However, the information about the axis of rotation is lost in this notation. In sequence development this axis can be of major importance when applying excitation pulses. Therefore, pulses are often described by their flip angle together with the rotation axis, e. g. 90°_x for a 90° pulse around the x -axis.

Physically, the excitation of magnetization is achieved using a radio-frequency (RF) coil, which is resonant to the desired frequency ω_0 [MLB15]. The received signal is generated by the same principal considerations as before. The rotating magnetization vector creates a magnetic field, which induces a voltage in the coil proportional to the transverse magnetization M_{\perp} . This process is governed by the well-known Faraday's law of induction:

$$\int_{d\Omega} \vec{E} \, d\vec{l} = - \iint_{\Omega} \frac{d\vec{B}}{dt} \, d\vec{A} \quad . \tag{1.12}$$

1.4 RELAXATION

In equation (1.9), \vec{M} precesses around B_0 indefinitely once it is excited (i. e., if $\vec{M} \nparallel \vec{B}_0$). In an experiment it is observed that the magnetization vector returns to equilibrium via a relaxation mechanism.

The T_1 , or spin-lattice relaxation, describes the energy transfer of spins with their surrounding (lattice). Nearby spins can couple to each other and

transfer energy by a mechanism involving the dipole-dipole interaction and the tumbling motion of the molecules. The relaxation causes the spin system to return to equilibrium \vec{M}_0 at an exponential rate, which is defined by the longitudinal relaxation time T_1 :

$$M_{\parallel}(t) = M_{\parallel}(0)e^{-t/T_1} + M_0 \left(1 - e^{-t/T_1}\right) \quad . \quad (1.13)$$

The T_2 , or spin-spin relaxation, describes the loss of coherence in a sample of spins. In any finite volume, the microstructure of the magnetic field is a superposition of the homogeneous external field B_0 and additional randomly fluctuating fields B_{rand} arising from the molecular fine structure. Since the Larmor frequency is dependent on the actual magnetic field at the position of the particle, different spins will have different Larmor frequencies, and, therefore, lose coherence (i. e. lose their constant phase relationship). This results in a decay of the transverse magnetization with the transverse relaxation time T_2 :

$$M_{\perp}(t) = M_{\perp}(0)e^{-t/T_2} \quad . \quad (1.14)$$

The relaxation times T_1 and T_2 can be added to equation (1.9) to formulate the *Bloch equations*:

$$\begin{aligned} \frac{dM_x}{dt} &= \gamma \left(\vec{M}(t) \times \vec{B}(t) \right)_x - \frac{M_x(t)}{T_2} \\ \frac{dM_y}{dt} &= \gamma \left(\vec{M}(t) \times \vec{B}(t) \right)_y - \frac{M_y(t)}{T_2} \\ \frac{dM_z}{dt} &= \gamma \left(\vec{M}(t) \times \vec{B}(t) \right)_z - \frac{M_z(t) - M_0}{T_1} \quad . \end{aligned} \quad (1.15)$$

In equation (1.15), the magnetization vector is precessing around B_0 with an exponentially decreasing amplitude of the transversal component and a parallel component that tends toward B_0 . Figure 1.2 shows the dynamics of an initial state $M(t=0)$ to a later time t .

For MRI, it is important that T_1 and T_2 are not constant for different tissue types. Thus, different tissue contrasts can be achieved by designing experiments with different weightings on T_1 , T_2 , and the spin density ρ (equation (1.8)).

1.5 RESONANCE EFFECTS - THE ECHO

A fundamental concept in MRI is the spin echo. For a thorough description of an echo in MRI, the review [hennigEchoesHowGenerate1991] is an excellent

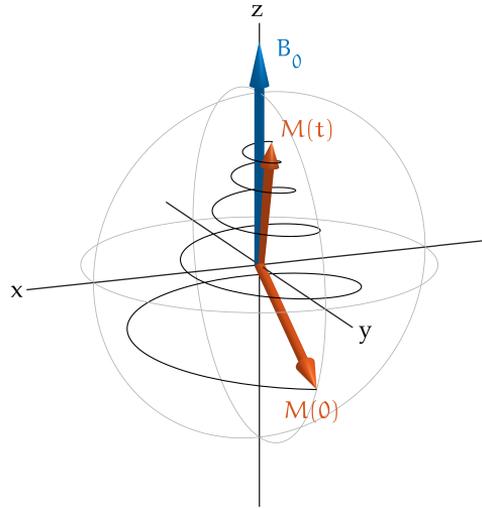


Figure 1.2: Dynamics of M , which is not in equilibrium state at time $t = 0$. In the presence of just the B_0 field, M precesses around the z -axis while it loses coherence in the x - y -plane (reduced amplitude, T_2 -decay) and simultaneously tends toward its equilibrium state along B_0 (T_1 -decay).

source. Here, a short summary of the most important types of echoes will be presented.

Consider the case where a magnetization in equilibrium state is excited by a 90° pulse. The resulting signal, measured by a nearby receive coil, will start at maximum amplitude and oscillate at a frequency ω_0 , while it decays with the time constant T_2 . This is called the *free induction decay* or FID. In MRI, FID experiments are not used, as they do not provide any spatial information. Instead, the signal is deliberately dephased during excitation, e. g. for slice selection (see chapter 2). The dephasing is achieved by applying an additional magnetic field, which varies linearly in space and is called the *gradient* G (unit: [T/m]), with

$$\vec{G} = \nabla B \equiv \text{const.} \quad . \quad (1.16)$$

Note, that the scalar field $B = B(x, y, z)$ is the z -component of the vector field \vec{B} . However, ignoring Maxwell-terms, the x - and y -components of \vec{B} are given by $B_x = B_y = 0$.

Assuming that a gradient G_z is applied in z -direction, the total magnetic field has the form

$$B(z) = B_0 + G_z \cdot z$$

$$\text{with } G_z = \frac{dB}{dz} \quad . \quad (1.17)$$

Thus, the Larmor frequency ω is spatially varying:

$$\omega(z) = \omega_0 + \Delta\omega = \omega_0 + \gamma G_z \cdot z \quad . \quad (1.18)$$

The spatially varying field leads to a signal dephasing, as the spatial phase offset $\Delta\phi$ is given by

$$\Delta\phi(z) = \int \Delta\omega(z) dt = \gamma \int G_z \cdot z dt \quad . \quad (1.19)$$

This dephasing is similar to the T_2 decay, but is reversible due to its deterministic nature: if the negative of the 0th gradient moment $\int G_z dt$ is applied later, the spins will rephase ($\Delta\phi = 0$) and the signal will be restored (i. e., only the contribution of G_z ; T_2 decay can not be compensated). For an arbitrary gradient $\vec{G}(t)$, the signal maximum is now observed at a later time TE, when the total gradient moment becomes zero:

$$0 \stackrel{!}{=} \gamma \int_0^{TE} \vec{G}(t) dt \quad (1.20)$$

This is called a *gradient echo*, and the basic principle is shown in figure 1.3 [Lau73].

In tissue, additional, inhomogeneous magnetic field gradients are present that are caused by susceptibility differences. In contrast to the random fluctuations causing the T_2 decay, these gradients are stationary. They lead to an additional signal decay mechanism T_2' , resulting in an effective decay T_2^* , with

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'} \quad . \quad (1.21)$$

The *spin echo* [Hah50] restores coherence of a spin isochromat (i. e., an ensemble of spins with identical Larmor frequency ω_0 in the absence of gradient fields), which has been dephased by the T_2' related gradients. Therefore, after a time TE/2 after free precession (e. g., after a 90° excitation pulse), a 180° *refocusing* pulse is applied to the sample. Since-susceptibility-induced gradients are stationary, the spin echo refocuses the T_2' component, resulting in a signal amplitude proportional to e^{-t/T_2} . The spin echo can be formed if the echo condition (equation (1.20)) is fulfilled, i. e. also without any gradients ($G(t) = 0$). The gradient echo, on the other hand, cannot discriminate between T_2 and T_2' , such that the amplitude is proportional to e^{-t/T_2^*} .

The spin echo can be understood by tracing the signals phase: Therefore, consider a constant gradient, which dephases the isochromat. The refocusing pulse inverts the phase accumulated during the time TE/2, such that after a

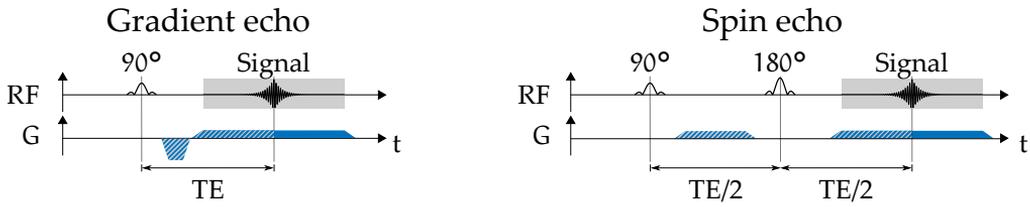


Figure 1.3: Schematic of a simple gradient echo and spin echo experiment. The echotime TE is the time after the excitation pulse (90°) when the spin coherence is restored. The signal amplitudes are proportional to e^{-t/T_2^*} for the gradient echo and e^{-t/T_2} for the spin echo. The striped regions are of equal area.

total time of TE the net accumulated phase of all spins is zero. Mathematically, a spin accumulates a phase ϕ_1 from the first gradient \vec{G}_1 , with

$$\phi_1 = \gamma \int_0^{TE/2} \vec{G}_1(\vec{x}, t) \cdot \vec{x} dt \quad . \quad (1.22)$$

The excitation pulse inverts ϕ_1 and after time TE the total phase is then given by

$$\phi_2 = -\phi_1 + \gamma \int_{TE/2}^{TE} \vec{G}_2(\vec{x}, t) \cdot \vec{x} dt \quad . \quad (1.23)$$

If

$$\int_{TE/2}^{TE} \vec{G}_1(\vec{x}, t) \cdot \vec{x} dt = \int_{TE/2}^{TE} \vec{G}_2(\vec{x}, t) \cdot \vec{x} dt \quad (1.24)$$

is fulfilled, the total phase ϕ_2 at TE is 0 and an echo is observed. Since the susceptibility induced gradients are stationary, this condition is always met. A simple spin echo example is shown in figure 1.3. Note, that the same echo appears even if the refocusing pulse is applied with a different phase than the excitation pulse, e. g. at a 90° angle from the excitation. Still, after the time TE, all spins will be in phase again, but the angle ϕ_2 can be unequal to 0.

Apart from the gradient echo and spin echo, higher order spin echoes and *stimulated echoes* can be created by applying multiple pulses after signal excitation. A more general description of the echo formation can be derived for arbitrary pulses with flip angle α . Each pulse can then be split into three components: a pure 0° , 90° , and 180° component with different weightings. Thus, each pulse acts partially as an excitation pulse and a refocusing pulse simultaneously. Figure 1.4 shows four different kinds of signal formations, given these considerations: b) shows primary (or first order) spin echoes, resulting from the first excitation pulse. c) shows a second order spin echo, which is created by the 90°

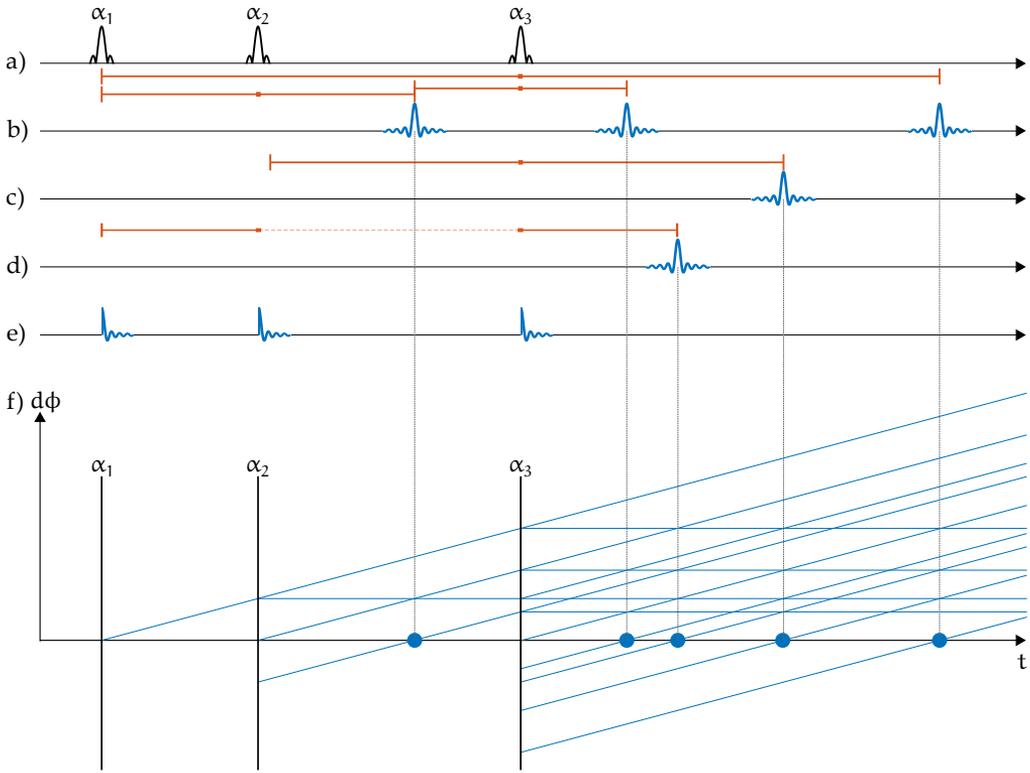


Figure 1.4: Graphical representation echoes: a) Excitation pulses. b) Spin echoes. c) Second-order spin echo. d) Stimulated echo. e) FIDs. f) Corresponding phase graph diagram. Figure adapted from [BKZ04, p.310, fig. 10.9] and concepts from [hennigEchoesHowGenerate1991].

component of the second pulse in combination with the 180° component of the third pulse. d) shows a stimulated echo, which is a result of the combination of all three pulses. Graph e) shows the FIDs created by each 90° component.

Finally, figure 1.4 f) shows the corresponding *extended phase graph*, which can be used to track the echo formation for any series of pulses. The blue lines track the phase evolution of an isochromat, splitting any line before the pulse into four new parts:

- one part continuous without being affected (0° component),
- one part is stored on the z-axis, i. e. the phase stays constant,
- one part is newly excited, i. e. a new line starts at $d\phi = 0$,
- one part continues with inverted phase, i. e. $d\phi_{\text{post}} = -d\phi_{\text{pre}}$.

At the zero-crossings (marked by blue dots), $d\phi = 0$ for any isochromat and, thus, an echo is observed.

IMAGING

In the previous chapter, the principles of signal formation in MRI were described. Now, the systematic manipulation of the signal for spatial encoding will be explained (section 2.1). Two pulse sequences, the *Fast Spin Echo* (FSE) and the *Echo Planar Imaging* (EPI), are presented (section 2.2), and the principles of image acceleration are described (section 2.3). This chapter is built on the textbooks by Bernstein [BKZ04] and Levitt [Levo8], where a more detailed description of the imaging process and pulse sequence design can be found.

2.1 K-SPACE - MEASURING IN THE FOURIER DOMAIN

In section 1.5 magnetic field gradients were introduced to generate gradient echoes. The same principle of changing the Larmor frequency with respect to the position of the magnetization will now be used for the spatial encoding of an MR signal. In Cartesian MRI, there are three different types of spatial encoding gradients: The *slice selection gradient* G_s , which selects a tomographic slice in an object. The *frequency encoding gradient* G_x and the *phase encoding gradient* G_y , which, together, encode the two-dimensional image in the slice.

As shown in equation (1.18), the Larmor frequency has a spatial dependency during slice selection:

$$\omega(z) = \omega_0 + \gamma G_s \cdot z \quad . \quad (2.1)$$

In the previous chapter, RF-pulses resonant to ω_0 were used to excite the signal. However, if the RF-pulse is designed such that it covers a certain range of frequencies $\omega_0 \pm \Delta\omega$ and has no frequency components outside of this range, the pulse can selectively excite spins based on their spatial position and the gradient G_z (see figure 2.1).

After this slice selective RF excitation, transverse magnetization is only present in the excited slice, which can be treated like a two-dimensional distribution of the signal $S(x, y)$. The measured signal in the receive coil is given by

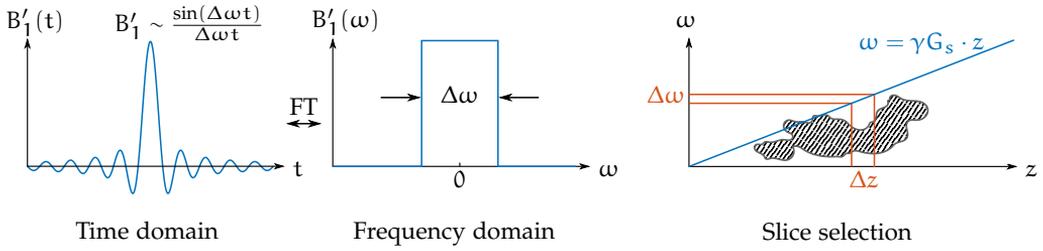


Figure 2.1: Principle of slice selection: In the time domain, the RF-pulse has the shape of a *sinc*-function, which translates to a rectangular bandwidth in frequency domain. Therefore, a rectangular region along the z -direction of the object can be selected.

the integral over all spins in the (2D) object, which are now modulated by the frequency (or readout) encoding gradient G_x , applied in x direction:

$$\begin{aligned} \tilde{S}(k_x) &= \iint S(x, y) e^{-ik_x(t)x} dx dy, \\ \text{with } k_x &= \gamma \int_0^t G_x(t') dt' \quad , \end{aligned} \quad (2.2)$$

where $S(x, y)$ is given by the spin density $\rho(x, y)$ and the relaxation times T_1 , T_2 and T_2^* . Note, that $\tilde{S}(k_x)$ has the form of a Fourier transform of $S(x, y)$, projected onto the y -axis.

Finally, *phase encoding* is applied prior to the frequency encoding. Therefore, consider the effect of the phase encoding gradient G_y along the y -axis. The gradient is turned off before the frequency encoding is turned on, and, thus, the signal $\tilde{S}(\vec{k})$ contains an additional, spatially dependent, phase:

$$\begin{aligned} \tilde{S}(\vec{k}) &= \tilde{S}(k_x, k_y) = \iint S(x, y) e^{-ik_x(t)x} e^{-ik_y y} dx dy, \\ \text{with } k_y &= \gamma \int G_y(t') dt' \quad . \end{aligned} \quad (2.3)$$

Note, that k_y is independent of t , because G_y is applied before the readout begins. To sample the complete image, the experiment is repeated N_y times with different gradient moments k_y in each experiment, where N_y is the desired matrix size.

In MRI, the space defined by k_x and k_y is called *k-space*, and the spatial coordinates x and y define the *image space*. As seen from equation (2.3), the acquired data is the superposition of all individual frequency components, encoded by read and phase gradients. This can be transformed into a spatial decomposition via the Fourier transform. An example of the relationship between k -space and image space is shown in figure 2.2.

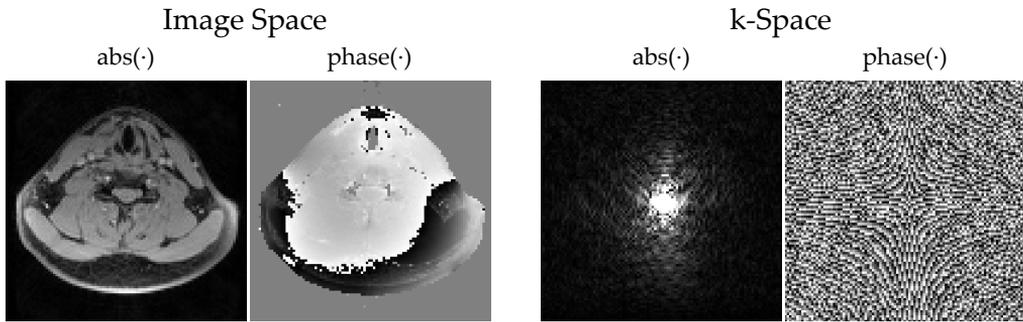


Figure 2.2: The human neck in a T_2 -weighted MRI. The left images display the magnitude and phase in image space, while the right side shows the same in k-space. Both representations contain the same information.

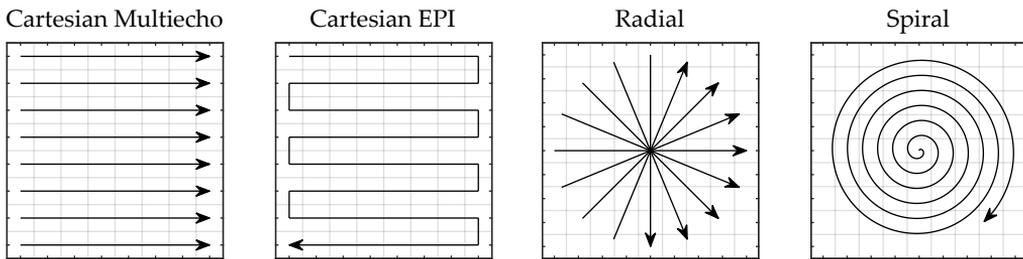


Figure 2.3: Different schemes k-space sampling. Often, k-space is sampled on a Cartesian grid (a) and b)), but generally arbitrary sampling patterns can be used. To reconstruct an image, which can be displayed on a computer screen, the image space representation has to be on a Cartesian grid.

The line-by-line acquisition method in k-space describes a Cartesian MRI, as shown in the first panel of figure 2.3. However, from equation (2.3) it can be seen that there are other, sometimes more efficient, ways to sample k-space. In echo planar imaging (EPI, panel 2), the complete k-space is sampled within a single excitation. If the readout is performed while the different gradients are switched on simultaneously, the effective gradient direction can be rotated, resulting in, e. g., radial or spiral k-space trajectories (panels 3 and 4).

It is important to note that the Fourier relationship between k-space and image-space is continuous. The measurement process however discretizes the acquisition matrix, so that the image can be obtained by a discrete Fourier transform. To be able to display the image, it has to be sampled on a Cartesian grid, but, as described before, the measured data can be sampled along an arbitrary trajectory. This discrepancy is usually solved by *gridding*, i. e. the interpolation of the measured data onto a Cartesian grid in k-space. From there, the Fast Fourier Transform (FFT) algorithm can be used to obtain the image.

2.1.1 IMAGE SPACE AND K-SPACE

The encoding of spatial positions through the Larmor frequency implies that the measurement must be sensitive to a certain range of frequencies. This *receiver bandwidth* (not to be confused with the bandwidth of the RF pulse), is connected to the field of view (FOV) of the image via the gradient amplitude. Given a receiver bandwidth BW , the lowest and the highest measured frequencies are

$$\begin{aligned}\omega_{\min} &= \omega_0 - \frac{1}{2}BW \\ \omega_{\max} &= \omega_0 + \frac{1}{2}BW \quad .\end{aligned}\tag{2.4}$$

With equation (1.18) and a readout gradient with amplitude G_r , we find

$$\begin{aligned}\omega_{\min/\max} &= \omega_0 + \gamma G_x \cdot x_{\min/\max} \\ \Leftrightarrow x_{\min/\max} &= \mp \frac{BW}{2\gamma G_x} \quad .\end{aligned}\tag{2.5}$$

From equation (2.5), the FOV is given by

$$FOV = x_{\max} - x_{\min} = \frac{BW}{\gamma G_x} \quad .\tag{2.6}$$

Now, let the dwell time Δt be the time between two acquired samples in k-space. K-Space and image space are connected through the Fourier relationship:

$$FOV = \frac{1}{\Delta k} = \frac{1}{\gamma G_x \Delta t} \quad ,\tag{2.7}$$

assuming a constant gradient G_x and discrete k-space steps Δk . Combining this with equation (2.6), it follows

$$\Delta t = \frac{1}{\gamma G_x FOV} \stackrel{(2.6)}{=} \frac{1}{BW} \quad .\tag{2.8}$$

The bandwidth plays an important role for the geometric accuracy of an image. If the main magnetic field is not completely homogeneous ($B'_0 = B_0 + \Delta B(x, y, z)$), the shift of the Larmor frequency at any location leads to a shift in the measured position, which is related to the pixel bandwidth $BW_{Px} = BW/N$, where N is the total number of acquired samples, i. e. the number of frequency encoding steps. Thus, a shift of the Larmor frequency by BW_{Px} due to ΔB leads to a shift of one pixel in image space, or, more generally,

$$\Delta x_{Px} = \frac{\gamma \Delta B}{BW_{Px}} \quad ,\tag{2.9}$$

where Δx_{px} is the image shift in pixels.

Considering this, a very high bandwidth always appears to be favorable. However, according to equation (2.8), a large bandwidth leads to short dwell times. The dwell time on the other hand can be seen as a signal averaging time, and therefore affects the SNR. In general, a compromise between bandwidth and SNR has to be found to acquire distortion free, high SNR images.

2.2 PULSE SEQUENCES

An MRI pulse sequence is a sequence of RF-pulses, gradients in x -, y - and z -directions and data acquisition blocks. Pulse sequences are designed to probe tissue parameters, such as T_1 or T_2 , or to generate an image from a composition of these parameters (contrast-weighted image). The echo time TE (see section 1.5) and the repetition time TR are two key parameters of a sequence, which often determine the expected contrast. The repetition time TR is defined by the time between two excitation pulses and is closely related to the T_1 contrast.

Here, the Carr-Purcell-Meiboom-Gill (CPMG) and the EPI sequence are presented, which are closely related to the work in chapters 5 and 6.

2.2.1 CARR-PURCELL-MEIBOOM-GILL SEQUENCE

The first sequence, shown in figure 2.4, is based on the CPMG (Carr-Purcell-Meiboom-Gill) sequence [CP54] and known as RARE [HNF86]. Today, RARE is commonly called fast spin echo (FSE) or turbo spin echo (TSE), which are vendor-specific terms for the same sequence with only few modifications.

The CPMG sequence starts with a slice selective 90° pulse, which is, after time $\Delta TE/2$, followed by a train of slice selective 180° refocusing pulses to create and refocus spin echoes. Note, that often the refocusing pulses can be applied with $\alpha < 180^\circ$ in order to reduce the energy deposited in the body. Typically, a Cartesian readout with fully refocussed readout and phase encoding gradients (i. e., they have a gradient moment of zero) is employed. Additional gradients, called crusher gradients, around the slice selection pulses are used to suppress FID signals, which occur for refocusing pulses that do not perfectly invert the signal [BKZo4, pp.305 ff.]. Imperfect refocusing is especially relevant due to limitations of the RF slice-profile, which can only approximate a rectangular shape for finite times of the RF-pulse.

Specific to the CPMG sequence are two conditions:

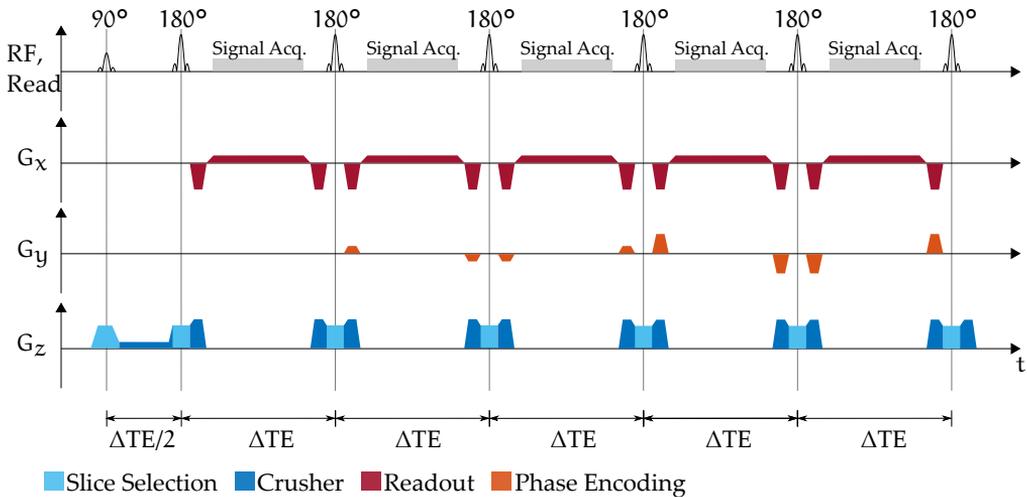


Figure 2.4: Sequence diagram for the RARE sequence. The signal is refocused multiple times, and each readout block is sorted in a different phase-encoding line of the same k-space. Here, the center of k-space corresponds to the first echo, and phase encoding is centric, i. e., positive and negative phase encoding steps are acquired in alternating echo polarity.

1. The refocusing pulses must have a phase shift of 90° with respect to the 90° pulse, and are spaced at an even time interval (ΔTE), which is twice the time interval between the excitation and first refocusing pulse ($\Delta TE/2$). Note, that other phase-cycling combinations than the 90° with respect to the excitation pulse can be found to achieve the same result.
2. The gradient moments between any two refocusing pulses are identical.

The first condition ensures artifact suppression from imperfect refocusing pulses and B_1 field inhomogeneity. The second condition guarantees that higher order spin echoes and stimulated echoes coincide with the primary echo. In combination, the conditions create multiple echoes with identical phase that interfere constructively.

Typically, the number of refocusing pulses before the next excitation (echo train length ETL) ranges between 8-32 echoes. To acquire a fully sampled image, the sequence is repeated until the total number of phase encoding lines is reached, resulting in total image acquisition times of a few seconds for an entire slice.

The echoes in the sequence are sampled at a different TE. Hennig et. al. [HNF86] proposed to sort all echoes into a single k-space, such that the shortest TE (the first echo) is placed in k-space center and echoes with longer TE are placed in the k-space periphery. Since the signal amplitude is proportional to

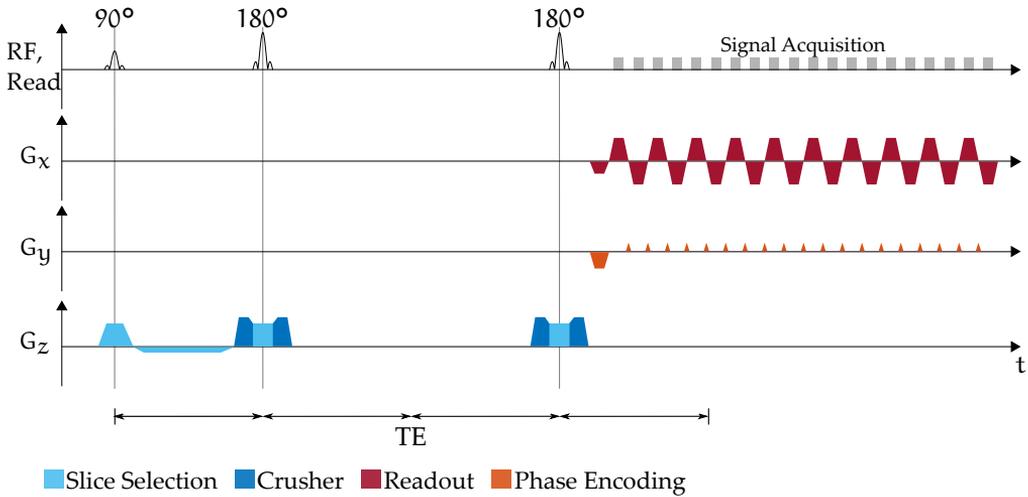


Figure 2.5: Sequence diagram of a double-refocused spin echo EPI sequence. During signal acquisition, the complete k-space is traversed in a meandering-fashion (see figure 2.3 b)). The phase encoding lines are changed by small blips in between the readouts, which switch direction every line.

e^{-t/T_2} , the signal decay acts as a filter in k-space, effectively blurring short T_2 components along phase encoding direction. Still, the image quality is sufficient for many clinical applications, and the gain in acquisition speed prevents problems such as patient motion during an exam. Note, that in order to achieve different T_2 contrasts, the central k-space line can also be acquired later in the echo train. Depending on the sorting of echoes into k-space, different T_2 filtering applies.

2.2.2 SPIN ECHO EPI

The spin echo EPI sequence is presented as a precursor for a diffusion sequence (section 3.2.2). The sequence can be split into the signal preparation (spin echo, SE) and the signal readout (EPI).

Figure 2.5 shows the sequence diagram of the SE-EPI sequence: The SE signal preparation includes a 90° excitation pulse and two 180° refocusing pulses. This setup is called *double-refocused* spin echo – one SE is formed after the first refocusing pulse at time $TE/2$ and refocused again at time TE . Slice selection gradients are applied as in 2.2.1, and crusher gradients again prevent signal interference from FIDs.

The EPI readout is a method to sample the complete k-space after a single excitation [STM91]. Therefore, a train of gradient echoes is acquired: the first

gradient echo is sampled in positive k_x direction, which is followed by a small step in phase encoding direction (*blip*), and another gradient echo in negative k_x direction. This is repeated until the complete k -space is traversed. The timing is chosen such that the spin echo coincides with the $k_y = 0$ gradient echo.

A main advantage SE-EPI sequence is the acquisition speed, which allows to sample a complete image slice in less than 100 ms [STM91].

2.3 ACCELERATED IMAGING

Acquisition speed is a major concern in pulse sequence design for several reasons: long acquisition times can lead to severe patient discomfort, resulting in incomplete examinations (premature interruption); patient motion is unavoidable for long acquisition times, leading to inconsistent data; or dynamic processes, e. g. heartbeat, breathing motion or systemic distribution of contrast agent, require a certain temporal resolution, which can only be met with fast sequences. Finally, faster image acquisition can also increase patient throughput, which is economically beneficial.

Here, image acceleration is split in two topics: first, hardware based acceleration using multiple receiver coils and, second, image acceleration by deliberate undersampling of acquired data.

2.3.1 PARALLEL IMAGING

So far, the MRI signal was assumed to be acquired by a single receive RF-coil, which covered the whole FOV. Parallel imaging is based on the application of multiple receive coils with limited FOV [Des12]. Importantly, each receive coil has a sensitivity profile, with higher sensitivity the closer to the coil [Roego; KEN91]. Consequently, the measured signal is a product of the coil sensitivity and the true signal (in image space).

Image acceleration is then achieved by deliberately skipping phase encoding lines during signal acquisition. According to equation (2.6), this undersampling of phase encoding lines leads to an increase in Δk , and, therefore, to a decrease in FOV. As a result of the Nyquist sampling theorem, the parts of the image that are now outside the FOV will be folded in and possibly overlap with the original image (wrap-around or aliasing artifact, see for example [Pus88]). Using a single receive coil, the overlapping signal cannot be separated, but with multiple coils with different spatial sensitivities, multiple unique images of the same object are acquired simultaneously. These images are formed by a linear combination of the true full-FOV image and the coil sensitivities. A

matrix inversion of this set of linear equations can then recover the complete image. Principally, this procedure can be performed in k-space as well as in image space, and the two corresponding algorithms are known as GRAPPA [Gri02] and SENSE [Pru99].

2.3.2 UNDERSAMPLING

Image acquisition can generally be accelerated by acquiring less data, i. e. reducing the number of echoes. In the previous section, this reduction of data was compensated by the information gained by additional receive coils. Alternatively, prior knowledge about the image or the image statistics can be used to recover the lost information:

First, *partial Fourier imaging* [Fei86] is based on the assumption that the image to be reconstructed is primarily real, i. e. its phase contribution does not contain the structural information of the image. For an entirely real image, its Fourier transform is Hermitian, i. e. all the image information is contained in one half of k-space. In practice, different amounts between $1/2$ and 1 times the k-space can be acquired, compromising between acceleration factor and SNR. Even if a phase component is present, different iterative algorithms can be used to recover the image [MPV88; McG93; HLL91].

Second, *compressed sensing* is based on the assumption that there is a sparsity transform, i. e., a transformation of the image into a space where the majority of the information content can be represented by only a few basis vectors [Dono6; BUf07]. For MRI, this relation is often fulfilled because within certain areas such as organs, blood vessels, gray or white matter in the brain, the signal variation is very small. Therefore, a transformation into wavelet space [Mey93] or application of the total variation (TV) [ROF92] yields only few non-zero coefficients (noiseless case). This property is exploited by choosing undersampling patterns, which produce noise-like artifacts in the image after Fourier transform. Then, the image x is reconstructed by an iterative algorithm minimizing a *loss function* Φ , which takes into account two parts:

1. Data consistency: $\|\mathcal{F}(x) - y\|_2 \stackrel{!}{=} 0$, and
2. Regularization: $\|W(x)\|_1 \rightarrow 0$.

Here, y are the acquired data, $\mathcal{F}(\cdot)$ is the Fourier transform and $W(\cdot)$ is the wavelet or TV transform. Together, the compressed sensing algorithm minimizes the following function:

$$\Phi = \arg \min_x \left(\|\mathcal{F}(x) - y\|_2 + \lambda \|W(x)\|_1 \right) . \quad (2.10)$$

The additional parameter λ is a weighting parameter between data consistency and regularization, which is usually determined empirically.

Compressed sensing is especially useful for sequences with non-Cartesian trajectories, where conventional parallel imaging or partial Fourier methods are not applicable. In radial imaging, fewer spokes can be acquired, such that the outer k-space areas are highly undersampled, while the k-space center is fully sampled (or even oversampled). This is often combined with an angular increase of the spokes by the golden angle, which increases the artifact incoherency necessary for the regularization term[Win07].

QUANTITATIVE MRI

The main clinical application of MRI is anatomical imaging with high soft tissue contrast. Depending on the anatomical region and the medical indication, T_1 -weighted, T_2 -weighted or spin-density weighted images are required. Especially in cancer imaging protocols, diffusion weighed imaging and dynamic contrast enhanced (perfusion) imaging are additionally included in standard protocols [ST16; KNo8].

With the exception of DWI, quantitative imaging is rarely used in clinical MRI. However, recent developments show that the quantitative parameter T_2^* could be used as a surrogate parameter for the detection of hypoxia in squamous cell carcinoma of the head and neck (HNSCC), where positron emission tomography (PET) is the current gold standard, often using ($[^{18}\text{F}]$ -fluoromisonidazole FMISO) tracer molecules [Wie18].

Additionally, with the increasing usage of AI in clinical decision-making, imaging protocols can be optimized for computer vision. As the parameter maps are the basis of any contrast weighted MR image, they might carry a larger information content than the contrast weighted images [Bie20b]. In addition, parameter maps are independent of the specific MRI system (assuming the same field strength), imaging protocol, or vendor, which is typically not the case for conventional imaging protocols due to different sequence implementations. Since most AI methods are dependent on large amounts of data to perform well, this intrinsic homogenization of quantitative MRI data could prove to be beneficial in future studies.

The following sections describe methods for T_2 and apparent diffusion coefficient (ADC) measurements, which are at the basis for the experiments in chapters 5 and 6.

3.1 T_2 MEASUREMENTS

In T_2 measurements (1.5), the spin echo recovers the loss of coherence due to the T_2^* -effect, and the SE formed by the refocusing pulse has only experienced a

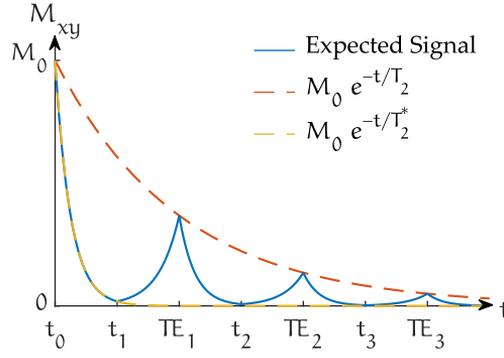


Figure 3.1: Envelope of the signal magnitude for a multi-echo-sequence. The expected signal (blue curve) is the result of the signal subjected to T_2^* decay (yellow line), and the refocusing of signal due to the 180° pulses which recover the signal up to an equivalent of pure T_2 decay (orange line). The refocusing pulses are played at times t_1 , t_2 , t_3 , and echoes occur at times TE_1 , TE_2 , TE_3 .

T_2 decay. Repeated application of refocusing pulses, as it is done in the CPMG sequence, will then repeatedly restore the coherence. Figure 3.1 shows the signal behavior.

For T_2 parameter mapping, the signal from each echo is assigned to a unique k-space, with a specific echotime TE_i . The sequence is repeated for all phase encoding steps, until the k-space for each TE is fully sampled. The contrast of each image is given by:

$$S_i = S(TE_i) = S_0 e^{-TE_i/T_2} \quad , \quad i = 1, \dots, n \quad (3.1)$$

where S_0 is the signal at $TE=0$ and n is the number of unique TEs. The T_2 parameter map can be obtained by a pixelwise linear regression of the following equation:

$$\begin{pmatrix} -TE_1 & 1 \\ -TE_2 & 1 \\ \vdots & \vdots \\ -TE_n & 1 \end{pmatrix} \cdot \begin{pmatrix} (T_2)^{-1} \\ \ln(S_0) \end{pmatrix} = \begin{pmatrix} \ln(S(TE_1)) \\ \ln(S(TE_2)) \\ \vdots \\ \ln(S(TE_n)) \end{pmatrix} \quad . \quad (3.2)$$

This general concept can be applied to any series of SE images with different TE, if the remaining signal contributions, e. g. T_1 contrast, are identical for each image in the sequence.

3.2 DIFFUSION

Molecular self diffusion describes the random movement of particles in liquids or gases. In the case of free diffusion, no boundaries restrict the movement, whereas in structured media like human tissue, particle diffusion is typically at least partly restricted. For example in neural axons in the brain, the diffusion along the fiber is essentially unrestricted, but across the fiber, there are tight cell boundaries which prevent particles from crossing.

MRI can be used to measure the diffusion without external tracers (e. g. dyes or pollen [Eino5]), since the MR signal stems from the tissue water itself. The basics of diffusion, and especially the theory of diffusion measurements using MRI, is excellently summarized in [Jon10], which also forms the basis for this section.

3.2.1 MOLECULAR SELF-DIFFUSION

The mathematical description of diffusion is based on Ficks first and second law, describing a net flux \vec{J} due to a gradient in particle concentrations n [Fic55]:

$$\begin{aligned}\vec{J} &= -D\nabla n(\vec{r}, t) \\ \frac{\partial n}{\partial t} &= D\nabla^2 n \quad .\end{aligned}\tag{3.3}$$

Treating the concentration $n(\vec{r}, t)$ as a local particle probability density, the diffusion coefficient D for a medium of spherical particles was derived (Sutherland-Einstein relation [Suto4; Suto5; Eino5]):

$$D = \frac{k_B T}{6\pi\eta R} \quad ,\tag{3.4}$$

where η is the viscosity of the medium and R is the radius of the particle in the liquid.

In general, the diffusion coefficient is then defined based on the autocorrelation function of the molecular velocity $v(t)$:

$$D = \lim_{t \rightarrow \infty} \int_0^t \langle v(\tau)v(0) \rangle d\tau\tag{3.5}$$

Since the molecular velocity can be anisotropic in structured media, D can be written in tensor form \hat{D} with entries

$$\hat{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix} \quad ,\tag{3.6}$$

forming an ellipsoid with different diffusivities in all spatial dimensions. Hence, there is an orthonormal coordinate system in which \hat{D} can be written in diagonal form. In general, however, the orientation of this coordinate system is unknown, and therefore \hat{D} is given by six independent values.

3.2.2 DIFFUSION MEASUREMENTS

To account for molecular self-diffusion in MRI, the Bloch-equations (equation (1.15)) were extended (Bloch-Torrey equations [Tor56]):

$$\begin{aligned}\frac{dM_x}{dt} &= \gamma \left(\vec{M}(t) \times \vec{B}(t) \right)_x - \frac{M_x(t)}{T_2} + \nabla \cdot D \nabla (M_x - M_{0,x}) \\ \frac{dM_y}{dt} &= \gamma \left(\vec{M}(t) \times \vec{B}(t) \right)_y - \frac{M_y(t)}{T_2} + \nabla \cdot D \nabla (M_y - M_{0,y}) \\ \frac{dM_z}{dt} &= \gamma \left(\vec{M}(t) \times \vec{B}(t) \right)_z - \frac{M_z(t) - M_{0,z}}{T_1} + \nabla \cdot D \nabla (M_z - M_{0,z}) \quad .\end{aligned}\tag{3.7}$$

Using this, the amplitude of measured signal additionally depends on the gradients as follows:

$$\begin{aligned}S(t) &= S(0) \exp \left(-D \gamma^2 \int_0^t \left[\int_0^{t'} G(t'') dt'' \right]^2 dt' \right) \\ &= S(0) \exp \left(-D \int_0^t \vec{k}(t') \cdot \vec{k}(t') dt' \right)\end{aligned}\tag{3.8}$$

The integral in equation (3.8) is known as the *b-value*:

$$b = \gamma^2 \int_0^t \left[\int_0^{t'} G(t'') dt'' \right]^2 dt' \quad .\tag{3.9}$$

Equation (3.9) shows that the b-value is approximately proportional to t^3 and G^2 . In practice, the short, low-amplitude gradients used in most imaging sequences will not add significantly to the b-value. For diffusion measurements, Stejskal and Tanner [ST65] proposed long rectangular *gradient pulse* pairs after excitation, which were separated by a 180° refocusing pulse. Then, the b-value can be simplified to

$$b_{\text{rect}} = \gamma^2 G^2 \delta^2 (\Delta - \delta/3) \quad ,\tag{3.10}$$

with G the gradient amplitude, δ the gradient duration and Δ the time between the two rectangular diffusion gradients. For more realistic trapezoidal gradients, this equation becomes

$$b_{\text{trap}} = \gamma^2 G^2 (\delta^2 (\Delta - \delta/3) + \zeta^3/30 - \delta \zeta^2/6) \quad ,\tag{3.11}$$

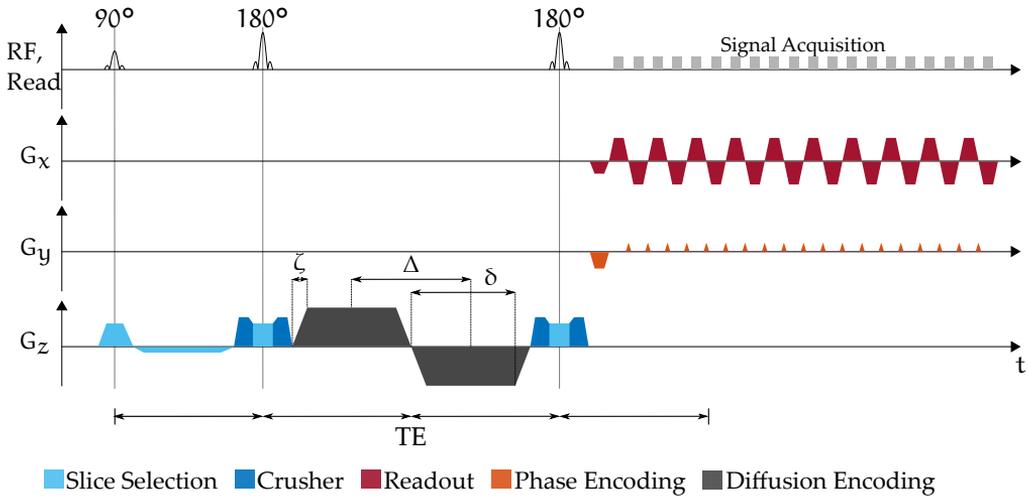


Figure 3.2: Schematic of a bipolar diffusion weighted EPI sequence. This sequence diagram is the natural extension of the sequence in figure 2.5 to incorporate diffusion weighing.

where ζ is the gradient rise, Δ the time between the gradient pulses and $\delta = \Delta - \zeta$ is the gradient flat time plus ζ [BKZo4, p.278]. The diffusion gradient can be rewound if instead of the 180° RF-pulse the second gradient has opposite polarity with respect to the first one, i. e. the total gradient moment after diffusion sensitization is zero.

The EPI-sequence presented in section 2.2.2 (figure 2.5), can now be extended into a diffusion weighted sequence by adding diffusion encoding gradients in between the two 180° pulses, as shown in figure 3.2. The b-value is then given by equation (3.11), and $\Delta = \delta + \zeta$.

Using this method, two measurements are sufficient to quantify the diffusion coefficient in an isotropic medium. However, in anisotropic media, additional measurements have to be performed, with at least six directions of the gradient axis, as discussed in the previous section. In practice, for a proper quantification of the diffusion tensor, often more than 100 measurements in different directions are taken to minimize errors. In clinical practice, however, diffusion tensor imaging is of minor importance. Instead, the ADC [Le 86] is often used for clinical decision-making. The ADC represents a mean of diffusion over multiple compartments with different diffusion values, as well as pseudo-diffusion such as vascular blood flow. The ADC can be computed by measuring three orthogonal diffusion directions of \hat{D} , calculating $D_{\alpha\alpha}$ for each of the directions

α by linearly fitting the logarithmic signal, and finally taking the mean of these values:

$$\begin{aligned}
 S_i &= S_0 e^{-b_i D_{\alpha\alpha}} \\
 \Leftrightarrow \ln(S_i) &= \ln(S_0) - b_i D_{\alpha\alpha} \\
 \Rightarrow D_{\alpha\alpha} &= \vec{b}^{-1} \cdot \left(\ln(\vec{S}_0) - \ln(\vec{S}_i) \right) \\
 \text{ADC} &= \frac{1}{3} \sum_{\alpha=1}^3 D_{\alpha\alpha}
 \end{aligned} \tag{3.12}$$

Because the trace of a matrix is independent of the choice of basis, the ADC is calculated using three orthogonal b-value directions. Hence, the ADC is independent on patient position or choice of the b-value direction.

3.2.3 ARTIFACTS IN DW-EPI

The additional diffusion sensitization gradients can lead to a number of practical problems in the MRI experiment. The *random phase artifact* in consecutive measurements is one of the reasons for the popularity of EPI for diffusion MRI. However, EPI in general suffers from *Nyquist ghost artifacts* and *geometric distortion* of the images. Finally, because of the random phase artifact, biased ADC measurements due to signal averaging on magnitude data are another common source of error in DW-MRI.

Random Phase Artifact

In a DW-MRI experiment, in contrast to molecular self-diffusion, patient motion is directed. The isochromats from a moving region will experience a total gradient moment which is not 0, resulting in a residual phase in the image. Since the diffusion gradient amplitude is often maximized for the given system, even small (e. g., pulsation) and involuntary patient motion will result in an additional phase. However, since the motion is not predictable, the phase artifact in the image appears to be random, leading to inconsistent image phases for consecutive measurements. Consequently, conventional SE or gradient-echo sequences, which sample k-space with multiple excitations, are not suited for DWI. Thus, to circumvent this phase artifact, single shot sequences such as EPI are used for the data readout.

Nyquist Ghost Artifact

The cause of the Nyquist ghost artifact are gradient timing errors, which introduce a small delay between the intended gradient execution and the actual

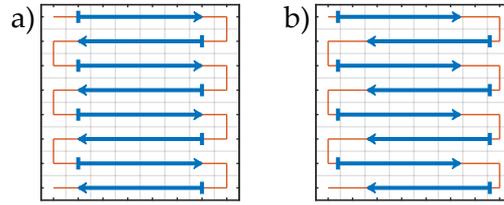


Figure 3.3: The k-space trajectory is plotted in orange, and the acquisition window is overlaid in blue. a) shows the k-space trajectory for an EPI measurement without gradient delay errors and b) shows the same with a gradient delay error. The echo always appears in k-space center ($k_x = 0$), but the acquisition window is not centered around this point in b), such that in the acquisition matrix each line appears shifted.

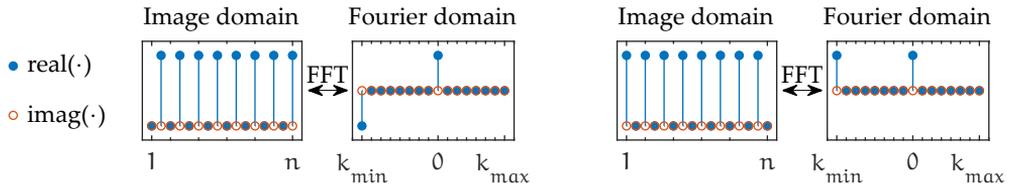


Figure 3.4: 2D image and Fourier domain representation of the mask which is applied to odd and even k-space for the simulation of the Nyquist ghost artifact.

onset of the imaging gradient. In k-space, this gradient delay results in a shift of each acquisition window along the read direction. Since the read direction is reversed in every other line, the actual k-space center ($k_x = 0$) is encoded at different positions k_x for even and odd k-space lines. This behavior is shown in figure 3.3.

For a mathematical description of the artifact, we can assume that k-space is separated into two parts by making use of the linearity of the Fourier transform:

$$\mathbf{k} = \mathbf{k}_{\text{odd}} + \mathbf{k}_{\text{even}} \quad , \quad (3.13)$$

where the odd k-space contains all odd lines and zeros instead of the even lines, and vice versa for the even k-space. Each space can be seen as the complete k-space multiplied with a mask that sets every second line to zero. The Fourier transform of the \mathbf{k}_{odd} -mask shows a delta-peak at $\mathbf{k} = 0$ and another one at $\mathbf{k} = \mathbf{k}_{\text{min}}$, while the \mathbf{k}_{even} -mask shows the same peak at $\mathbf{k} = 0$ and the second peak with inverted sign (figure 3.4). The multiplication in Fourier domain equates to a convolution in image domain, such that the second peak in either \mathbf{k}_{odd} or \mathbf{k}_{even} produces a ghost image shifted by half the field of view. However, the inverted sign in the \mathbf{k}_{even} -mask leads to a phase-shift of the ghost image with respect to the \mathbf{k}_{odd} -mask. Therefore, if odd and even k-spaces are not shifted due to gradient delays, the ghosts (shifted images) cancel out and only the original (unshifted) image is left. If a shift in k-space is present, it leads

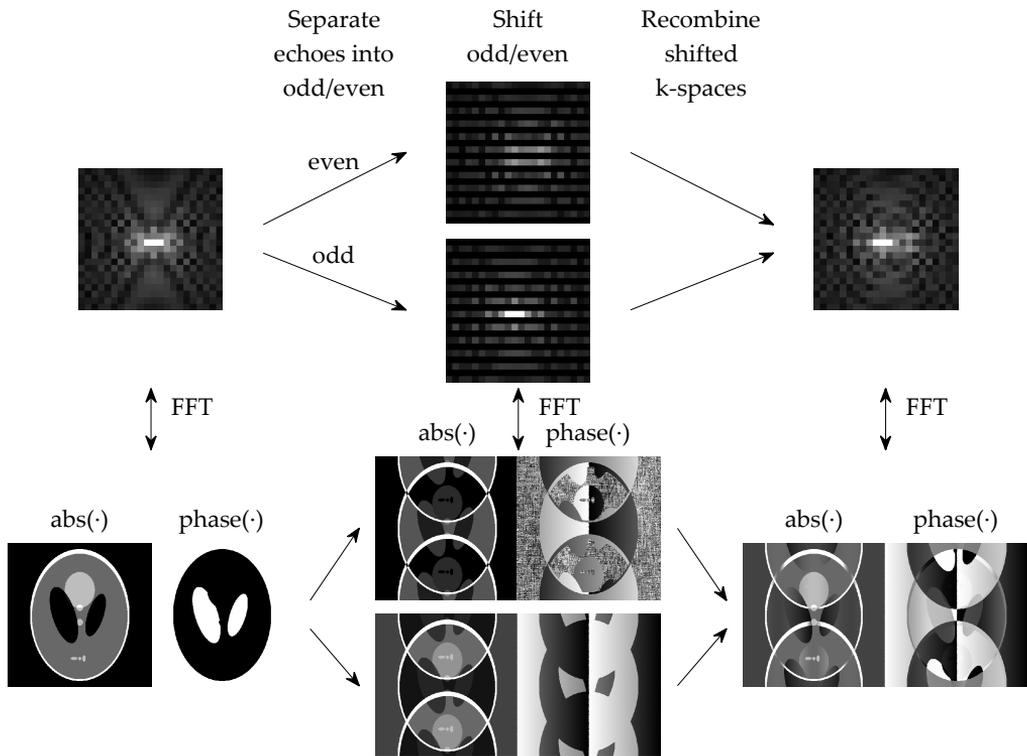


Figure 3.5: Simulation of the Nyquist ghost artifact in a EPI-acquired image. The true k-space (left) is shifted by $+1/-1$ pixel for even/odd k-space lines. Upon recombining the shifted k-spaces, the ghosts in image space sum incoherently, resulting in a residual ghost image.

to a linear phase in image space, according to the Fourier shift theorem. Hence, the ghosts do not sum coherently and cancel, but sum up incoherently resulting in the residual Nyquist ghost artifact. Figure 3.5 shows a simulation with a k-space shift of 1 pixel, both for the image and the Fourier domain.

If image acceleration methods are used, some k-space lines are not acquired. Therefore, the k-space is not divided into odd and even, but into additional sub-k-spaces where only every 3rd, 4th, 5th, etc. line contains data. For these higher acceleration factors, the above argument still holds, but the corresponding masks produce additional peaks in their respective Fourier domain, which are expressed as additional ghost images.

To prevent the Nyquist ghost artifact, the same k-space line can be acquired twice, with opposite readout direction, prior to the actual measurement. From the acquired data, the shift due to timing errors can be estimated by computing the cross-correlation of the two measured echoes. This shift can then be used to retrospectively correct the acquired EPI-data.

Geometric Distortion Artifact

In multi-echo Cartesian imaging techniques, each phase encoding line is subject to the same free precession period. In contrast, for EPI, the signal phase evolves freely and continuously between phase encoding lines. Therefore, an effective bandwidth in phase encoding direction is defined, which is

$$BW_{Ph} = \frac{1}{t_f} , \quad (3.14)$$

where t_f is the time needed for one frequency encoding line (i. e., the time taken by one of the red trapezoids in figure 2.5). BW_{Ph} is approximately a factor of N_x times lower than the frequency encoding bandwidth, where N_x is the number of samples per frequency encoding line. Therefore, B_0 offresonance leads to a significant shift of the corresponding signal in phase encoding direction, which manifests in image distortion as described in section 2.1.1. The amount of distortion is then given by

$$\Delta y_{Px} = \frac{\gamma \Delta B N_y}{BW_{Ph}} , \quad (3.15)$$

where N_y is the number of phase encoding steps and Δy_{Px} is the pixel-shift in phase encoding direction.

Signal Averaging of Magnitude Data

DW-EPI inherently suffers from low SNR for multiple reasons: First, the diffusion sensitization takes time, which results in relatively long TE of typically 50-100 ms, even if image acceleration is used. Because T_2 of soft tissues is in the same order of magnitude, a significant portion of the signal has decayed. Secondly, only in k-space center a SE is formed, such that all other lines are subject to T_2^* -decay, further decreasing the signal amplitude. Lastly, the diffusion-weighted signal relies on an exponential signal attenuation, such that for high b-values a considerable portion of the signal is deliberately destroyed.

For these reasons, signal averaging becomes crucial for data analysis. However, if the complex data were averaged directly, the random phase artifacts described in the beginning of this section would lead to signal cancellation in the image. To prevent those signal voids, only magnitude images are used for averaging.

The use of magnitude images has two major implications: First, the noise distribution in the magnitude image will be Rician [GP95]. The Rice distribution includes only positive values, is asymmetric for low SNR and converges against a Gaussian distribution for high SNR. The first property (only positive values)

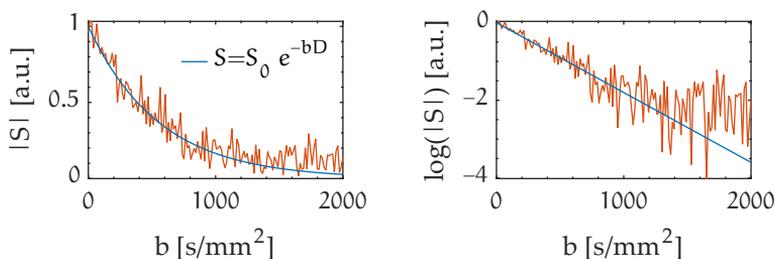


Figure 3.6: Ideal (blue) and noisy (orange) diffusion signal vs. b-value. The right plot shows the logarithmic signal, which is often used for ADC estimation. When the noisy signal approaches the noise level, a systematic deviation from the ideal signal becomes visible.

leads to a bias when several low SNR-images are averaged, since the noise does not add up strictly incoherently anymore. Figure 3.6 demonstrates this effect for a noisy diffusion signal. In the logarithmic scale, a clear deviation from the ideal exponential decay becomes visible for b-values larger than 1000 s/mm^2 . Since ADC maps are conventionally calculated by a linear regression of the logarithmic signal, a bias towards lower ADC is expected if low SNR images are used.

ARTIFICIAL INTELLIGENCE

The term artificial intelligence (AI) is used to describe an algorithm a machine (sometimes called agent) uses to perform a certain task; however, there is no coherent definition of the term among the scientific community [Wan19]. This work is centered around artificial neural networks (ANNs), which are one type of AI algorithm. The concepts of modern ANNs were invented in 1986 by Smolensky [RM86], but only in the last two decades these concepts could be effectively used. One milestone was the introduction of efficient training algorithms [Lec98; Hino2], and a second precondition for the training of modern neural networks (NNs) was reached due to the dramatic increase in computing power, especially with the rise of graphical processing units (GPUs) [Nil10]. An ANN *learns* to solve a problem by training on existing data, i. e., a model is established by training on data with known results. This is in contrast to model-based approaches, which assume a known relationship between the data and the result, and, therefore, no training is needed. This concept is also known as *learning from example*: the NN derives a functional relationship between input and output, similar to a parameter fit. However, the number of free parameters in the NN is very high (millions or even billions of free parameters), such that the network could principally model arbitrary functions [HSW89]. The quality of the network can then only be evaluated by its power to generalize to previously unseen data.

In this thesis, the focus lies entirely on pattern recognition, i. e. the automatic, pixelwise labeling of tumors in MRI images, given a large set of already labeled images. This kind of algorithm is characteristic for nearly all modern computer vision tasks, with ever increasing performances on image classification (i. e., assigning a label to an entire image) [Lec98; He15; KSH17] and segmentation (i. e., labeling to each pixel within an image) [Per15; RFB15; Kam17; Akk17; Hav17].

In the following sections, the concept of a convolutional neural network (CNN) is established. Thus, the necessary basic building blocks, the perceptron and the artificial neural network are introduced. Following this, the most

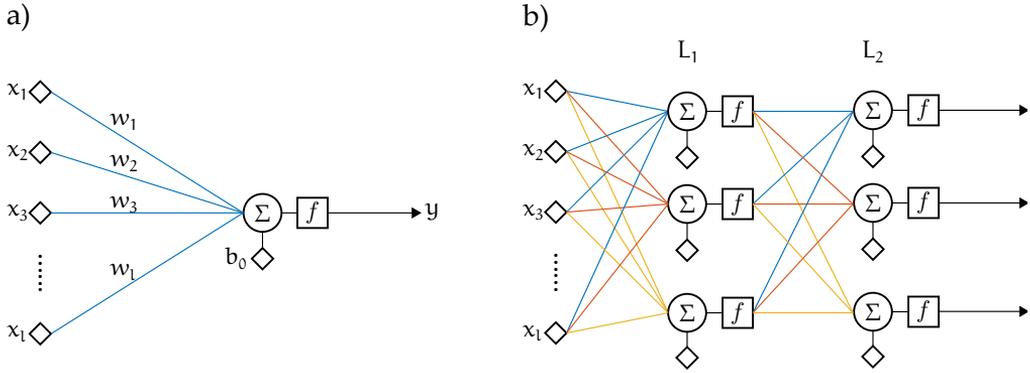


Figure 4.1: a) Schematics of a perceptron and b) a multilayer perceptron (MLP). Diamond shapes represent scalar values and each colored line represents a multiplication with a unique weight w . Each *node* consists of the summation (Σ) of the weighted inputs, the bias and a consecutive application of a (nonlinear) function f .

commonly encountered additional building blocks and the most important parameters of CNNs are presented.

The textbooks [TKo6] and [Nil10] are recommended sources for understanding the origins and algorithms involved in pattern recognition, and form the basis of the following sections.

4.1 THE PERCEPTRON

The *perceptron* was originally designed as a linear classifier. It takes scalar inputs x_i and assigns a weight w_i to each input. The output y of the perceptron is then given by

$$\begin{aligned} y &= b_0 + \sum_i w_i x_i \\ &= b_0 + \vec{w}^\top \vec{x} \quad , \end{aligned} \tag{4.1}$$

where b_0 is an additional scalar called *bias*, and $(\cdot)^\top$ is the transpose operation. This perceptron can be used to find a hyperplane (i.e. a multi-dimensional linear function) which separates inputs \vec{x} that belong to different classes ω_1 and ω_2 . This is achieved by finding an appropriate weight vector \vec{w} and bias b_0 , which yields $y > 0$ for all $\vec{x} \in \omega_1$, and $y < 0$ for the other. The algorithm to find the correct weights and bias is called the *perceptron algorithm*, and it

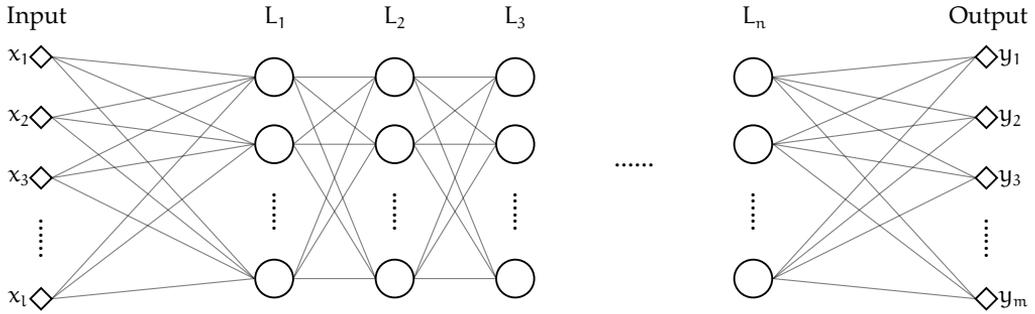


Figure 4.2: Schematic of a complete ANN. The nodes include the bias, the summation and the nonlinearity.

minimizes a loss function $J(\vec{w})$ using an iterative gradient descent method. The loss function is given by

$$J(\vec{w}) = \delta_x \mathbf{b}_0 + \sum_{\vec{x} \in X_{\text{mis}}} \delta_x \vec{w}^\top \vec{x} \quad , \quad (4.2)$$

and $\delta_x \equiv -1 \forall \vec{x} \in \omega_1 \quad ,$
 $\delta_x \equiv +1 \forall \vec{x} \in \omega_2 \quad ,$

where X_{mis} contain all vectors \vec{x} that were assigned to the wrong class y (misclassified).

The perceptron can be visually represented as shown in figure 4.1 a): Here, the additional function f is given by

$$f(x) = x \quad . \quad (4.3)$$

As shown later, f will be replaced by a nonlinear function for more complex cases.

4.2 ARTIFICIAL NEURAL NETWORKS

An artificial neural network is an assembly of multiple perceptrons, organized in a number of layers L_j (also known as multilayer perceptron MLP). In the example in figure 4.1 b), each input value x_i is connected to each *node* in layer L_1 with a weight w_{ij} . Hence, the output of the first layer are three values y_j^1 , given by

$$y_j^1 = f \left(b_j^1 + \sum_{i=1}^l w_{ij}^1 x_i \right) \quad . \quad (4.4)$$

Here, the superscript 1 is used to indicate the layer number. The output vector \vec{y}^1 is then the input vector for the next layer L_2 , following the same principle as in equation (4.4). To model nonlinear functions, the function f is introduced: It introduces a nonlinearity (e. g., a logistic sigmoid [Hin12]) which allows to model arbitrary functional relationships between input vectors \vec{x} and output vectors \vec{y} , given enough layers.

Figure 4.2 shows the complete schematic of an ANN. Here, the nodes, including the bias, the summation and the nonlinear function f are represented as a single circle. The input layer is given by \vec{x} , and the output layer by \vec{y} . The layers in between (L_1 to L_n) are called *hidden layers*. Each hidden layer can have a different number of nodes, which define the *width* of the layer. In the example shown here, the layers are referred to as *fully connected*, because each node is connected to every other node of the next layer. ANNs with more hidden layers are said to be *deeper* (hence the terms *deep neural net* and *deep learning*). Deeper networks show a higher capacity to model arbitrary functions due to their increasing number of nonlinear connections.

So far, the gradient descent algorithm to update the weights and biases in the network involved the minimization of a loss function based on all available data. Using this method, and assuming a convex loss function, a guaranteed minimum is found. However, with increasing network size and increasing amounts of input data, this problem becomes computationally too expensive, as the memory requirements cannot be met even by modern supercomputers. Thus, the successful training of a large ANN only became feasible by introducing the stochastic gradient descent (SGD) algorithm [RM51]: Intuitively, it states that the average gradient direction taken from calculating the gradients based on many single input vectors \vec{x} is the same as the gradient direction calculated from all \vec{x} at once. Hence, many iterations with small subsets of \vec{x} yields the same result as few iterations with all available \vec{x} . In order to achieve a satisfactory result, the randomly drawn input data \vec{x} may have to be used multiple times.

The process of iteratively adapting the learnable parameters (*learnables*) of the network (i. e., the weights and biases) is called *training*, and the calculation of the gradients given the input vectors \vec{x} is called *backpropagation*. Two terms are often encountered in the network training:

- The *epoch* is the training cycle when all \vec{x} have been seen once. Complete training of a network usually takes many epochs.
- The training *mini-batch* describes a compromise between pure SGD and conventional gradient descent. In SGD, all learnables are updated after a single input was processed. Using a mini-batch means that a larger number of inputs are processed simultaneously and the learnables are

updated based on the mean gradient over these inputs. However, the mini-batch can be much smaller than the complete training set, which would be the case for conventional gradient descent.

The specific implementation of the gradient descent optimization algorithm is independent of the network architecture. Many of these algorithms depend on a number of additional parameters (hyperparameters), but common to all of them is the *learning rate*, which controls the step size in the gradient direction when updating the weights. Often, an adaptive approach is chosen, which decreases the learning rate either after a preset amount of epochs, or when the performance change of the network during training becomes too small.

4.3 CONVOLUTIONAL NEURAL NETWORKS

In early attempts, image classification was performed by fully connected ANNs, using the complete image as an input vector, such that each pixel corresponds to one entry in \vec{x} . Hence, the input vector size scaled quadratically with the image size. Furthermore, the layers of fully connected networks typically contain at least as many nodes as the input, such that, for images, ANNs became computationally very expensive. For example, a 128×128 image with 5 fully connected layers of the same size results in a network that contains 1 342 259 200 learnable parameters. Since all of the parameters need to be updated in one training pass, they have to be stored in the computer memory. Hence, for larger networks, this method becomes unfeasible.

Convolutional neural networks (CNNs) are designed specifically for image processing. In a CNN, each input $x_i \in \vec{x}$ is an image (2D, 3D or arbitrary dimension) itself. The weights, previously just scalars, now become matrices that match the dimension of the input image, but can have a smaller size, e. g. a 3^n matrix, where n is the image dimension (2D, 3D, ...) (these weight matrices often are called *kernels*). In the ANN, there is a multiplicative relation between input and weight (equation (4.4)) – this relationship is replaced by a convolution with the kernel for CNNs. Therefore, equation (4.4) becomes

$$\hat{y}_j^L = f \left(b_j^L + \sum_{i=1}^l \hat{x}_i \otimes \hat{w}_{ij}^L \right) , \quad (4.5)$$

where \otimes is the convolution operator and the $\hat{(\cdot)}$ marks whether the variable is a matrix (or higher dimensional) object. The superscript L again marks the layer number.

A key difference of CNNs to fully connected ANNs is the decoupling of pixel position in an image and the associated weights. Common to all convolution

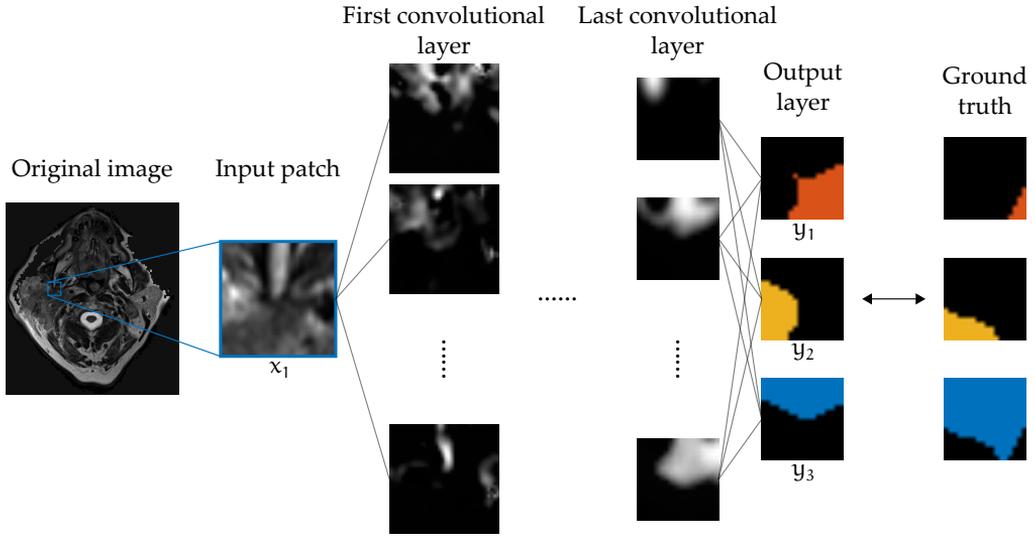


Figure 4.3: Example for a CNN. From a large image, a patch is extracted which is then processed in the CNN. In the convolutional layers, the feature maps as calculated by equation (4.6) are shown. In the end, the computed segmentation is compared against the ground truth to compute the gradients for network training.

kernels is the translational invariance, which is an inherent feature due to the convolution operation. Thus, structures are detected irrespective of their image position, which leads to a superior performance of CNNs in image classification over fully connected ANNs despite their reduced number of learnable parameters.

The convolution in a CNN is typically implemented as a sliding window multiplication. Mathematically, this can be described as follows: assuming 2D inputs \hat{x} and kernels \hat{w} , and a kernel size of $3 \times 3 = 3^2$, equation (4.5) can be written as

$$\begin{aligned}
 (\hat{y}_j^L)_{m,n} &= f \left(b_j^L + \sum_{i=1}^l (\hat{x}_i \otimes \hat{w}_{ij}^L)_{m,n} \right) \quad \text{and} \\
 (\hat{x} \otimes \hat{w})_{m,n} &= \sum_{u,v=-s_{\max}}^{s_{\max}} x_{m+u,n+v} \cdot w_{u,v} \quad .
 \end{aligned}
 \tag{4.6}$$

Here, (m, n) denotes the matrix indices, and $x_{m,n}$ and $w_{u,v}$ denote the respective matrix elements. In this notation, $w_{0,0}$ is the central element. Equation (4.6) is only defined for the ranges $1 \leq m \pm s_{\max} \leq m_{\max}$ and $1 \leq n \pm s_{\max} \leq n_{\max}$. Thus, the matrix size of the output \hat{y} is decreased, depending on the kernel size s_{\max} .

In CNNs, the intermediate outputs \hat{y}_j^L in each layer L are called *feature maps*. Figure 4.3 shows a CNN evaluated with an example input. From the original image, a *patch* of a smaller size is taken and subsequently processed in the convolutional layers. The kernels in the CNN can be seen as filters which identify recurring structures in an image. In early layers, they can often be interpreted as edge or intensity filters, while they become increasingly abstract in deeper layers.

In contrast to an ANN, a CNN cannot necessarily include all information from the image for the classification of a single pixel. The *receptive field* φ defines the neighborhood of pixels which contribute to an output pixel in a given feature map. Hence, in the last layer, it describes the amount of contextual information the CNN can utilize. The receptive field is calculated given the kernel sizes κ for each dimension and the layer number l :

$$\begin{aligned}\varphi_l &= \varphi_{l-1} + (\kappa_l - 1) \\ \varphi_{\text{last}} &= \sum_l (\kappa_l - 1) \quad .\end{aligned}\tag{4.7}$$

It can be shown that smaller kernel sizes (minimum of 3^n) can increase the networks' segmentation performance, because they achieve deeper networks with fewer learnable parameters but constant receptive field as compared to networks with larger kernel sizes [SZ15].

4.4 NETWORK ARCHITECTURE

A CNN is not only defined by the width and depth of the network: layers can include additional building blocks, and they can be arranged in more complex structures. Additionally, the choice of parameters within the network can have a significant influence on the performance. The choice of the best parameters is usually not clear a priori, and depends on the specific task and the data used to train the network. Unfortunately, this point applies to many of the design choices in a CNN, and the optimal architecture and parameter set can often only be found by excessive testing or systematic optimization procedures, like the Bayesian optimization [SLA12].

One of these important design decisions for a CNN is the activation function f . The initially proposed logistic sigmoid

$$f(x) = \frac{1}{1 + e^{-x}}\tag{4.8}$$

was the natural choice for the nonlinearity [Hino2]. However, when more efficient training algorithms were developed, f did not need to be differentiable

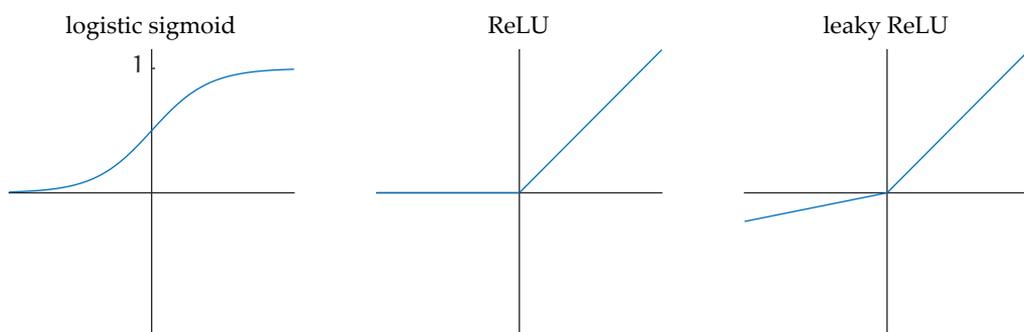


Figure 4.4: Three common activation functions for NNs. The logistic sigmoid is a shifted version of the tangent hyperbolicus. The leaky ReLU is linear ($y = x$) for $x \geq 0$ and 0 for $x < 0$. The leaky ReLU is an extension of the ReLU, where negative x are still linear, but with a small factor: $y = \alpha x, 0 < \alpha < 1, \forall x < 0$.

everywhere anymore, and functions that were only locally differentiable were proposed. The logistic sigmoid was especially unsuited for deep neural networks, where the neurons would saturate, which could be prevented by using other activation functions [GB10]. Two important activation functions are the rectified linear unit (ReLU) and the leaky ReLU (figure 4.4), but a plethora of others are available [Sza21].

The nonlinearity is a function specific to a convolutional or fully connected layer. However, in most networks additional layers are introduced, which serve different purposes:

- *(Batch-) Normalization* layers introduce additional learnable parameters, variance and a mean, which are used to normalize the data, and thereby increase the training speed [IS15]. The batch normalization layer can be applied after every hidden convolutional layer, and, thus, is different from a normalization during data preprocessing.
- *Pooling layers* are used to compress the network by decreasing the size of the feature maps, which can increase the networks performance [Nag11].
- *Residual connections* add the output of a previous layer to the output of the current layer. This operation can increase the training success, because the current layer is then trained to approximate the residuals of a function instead of the function itself, which can be beneficial in deep networks [He15].

The performance of a trained NN is not only dependent on the general architectural decisions, but also on the training process itself. The optimization algorithm can have a significant influence whether a minimum is reached

slowly or fast, or not at all [LeC12]. For example, one variant of the previously introduced SGD is called Adam, which updates the gradient direction based on a momentum rather than the the current gradient itself, which can help to reach the optimum more efficiently [KB17].

In addition to the optimization algorithm, the *cost function* plays a major role in the optimization process. The cost function (or loss function) is the function which is optimized by the optimization algorithm. In semantic segmentation, a typical example is the cross entropy, which computes the current cost based on the L2-norm, i. e., it sums over all misclassified pixels. Other cost functions implement the Dice Sørensen Coefficient (DSC)

$$\text{DSC} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad , \quad (4.9)$$

where TP are true positives, FP are false positives and FN are false negatives [Sud17], or variants of the cross entropy which modulate the loss based on contextual information like the focal loss [Lin17].

Often, the available training data for a specific task is limited or does not cover the complete space of expected data which the network should be able to generalize. To (partially) overcome this limitation, data *augmentation* is performed: The pool of available training data is extended by artificially modifying the existing data [PW17]. Typical data augmentation strategies involve rotation, flipping, or general affine transformations, as well as the artificial introduction of artifacts like noise or image blurring. The idea of data augmentation is to prevent *overfitting*, i. e. training the network to be highly accurate on the already seen training data, but performing poorly on unseen data. Another effective method to prevent overfitting is the introduction of additional *dropout layers*, which randomly deactivate connections in the network by either multiplying a weight (or kernel, in CNNs) with 0 or 1 [Sri14]. When the network training is finished, the dropout layers are deactivated, such that all weights always contribute to the final result.

Part II

MATERIALS AND METHODS

AUTOMATED TUMOR SEGMENTATION

5.1 FMISO STUDY

5.1.1 MRI DATA OF HEAD AND NECK TUMORS

In this thesis, the development of CNNs for tumor segmentation is centered on MRI data from the FMISO trial, a prospective clinical phase II trial [GN15]. The trial is designed to study the changes of head and neck squamous cell carcinomas under radiochemotherapy using PET-CT and MR imaging. Specifically, the response of hypoxic subareas within the tumor is investigated using the PET tracer [^{18}F]-fluoromisonidazole (FMISO). MRI is used to provide an accurate anatomical tumor delineation as well as additional functional contrasts, such as tissue perfusion, diffusion or T_2^* -parameter maps. In this context, possibilities for MRI based hypoxia imaging by a single or a combination of different MRI contrasts are investigated [Wie18]. Therefore, an extensive MRI protocol was designed, including 7 different MRI-contrasts. Figure 5.1 shows the central slice from an example dataset with all relevant contrasts. Further details on the measurement protocol are found in section 5.1.4.

The first patient was included in the FMISO study in August 2014, before the beginning of the works related to this thesis. After an extension to increase the total number of patients, recruitment is still ongoing. Within the scope of this thesis, 33 patients have been included, 24 of whom received the complete imaging protocol. According to the study protocol, patients received an MRI before treatment start (week 0), and at weeks 2 and 5 during radiation therapy. The therapy included 7 weeks of radiation treatment with a 2 Gy daily dose up to a total dose of 70 Gy. Additional concomitant chemotherapy with cisplatin (100 mg/m^2 body surface area) in weeks 1, 4 and 7 was administered. The trial was approved in advance by the local Independent Ethics Committee (reference no. 479/12) and was carried out in accordance with the Declaration of Helsinki (revised version of 2015). The trial was registered with the German Clinical Trial Register (DRKS00003830).

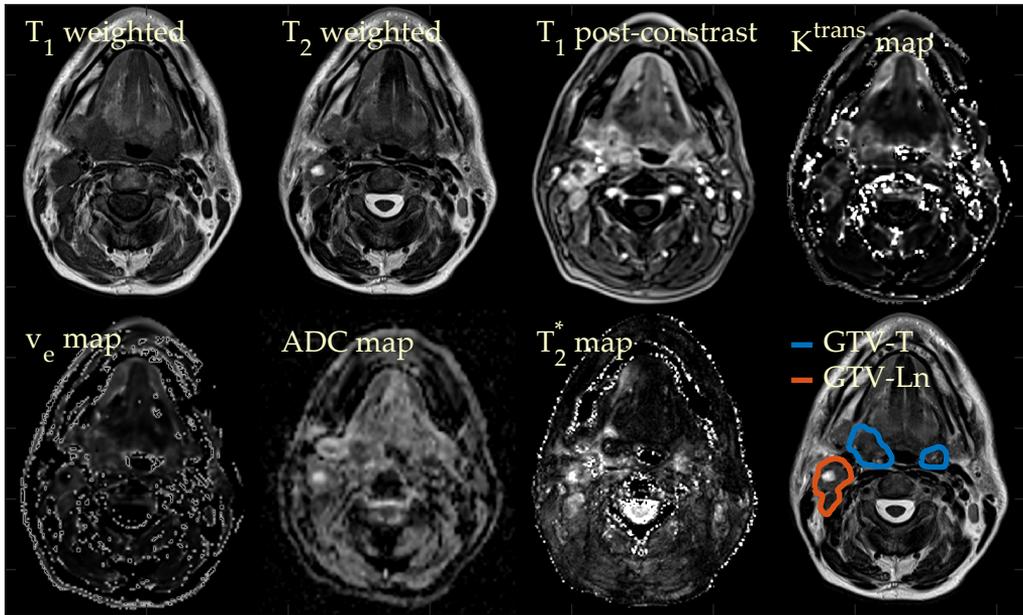


Figure 5.1: Seven different MRI contrasts of the same slice in a head and neck cancer patient. The last image (bottom right) shows the contours of the lymph node metastases (GTV-Ln) and gross tumor volume (GTV-T) overlaid on the T_2 -contrast. Figure adapted from [Biezob, fig. 1].

5.1.2 GROUND TRUTH SEGMENTATIONS

To provide a ground truth for the CNN training, manual segmentations were performed in the data set. Ground truth (GT) segmentations are logical masks which assign a label to each pixel in an image. The labels used in this work are background (BG), gross tumor volume (GTV-T) and lymph node metastases (GTV-Ln), typically associated with numbers 0, 1 and 2.

Since patients undergo radiotherapy, this segmentation must be of highest quality to avoid unnecessary damage to nearby healthy tissue without missing any of the diseased tissue. Hence, given the imaging data, GT tumor segmentations were generated in consensus by a radiologist and a radiation oncologist for each patient. This procedure was performed within the software environment *IPlan* (Brainlab AG, Munich, Germany). Within *IPlan*, the user imports MRI, CT and PET data and can display fused versions of the data. The underlying fusion algorithm is based on the mutual information metric [VI97] and can be applied semi-automatically, allowing manual adjustments based on sub-regions of the images. In the next step, the user defines different regions of interest, including GTV-T and GTV-Ln, by manually drawing the outline on a slice in one of the available image sets, while a co-registered version of the

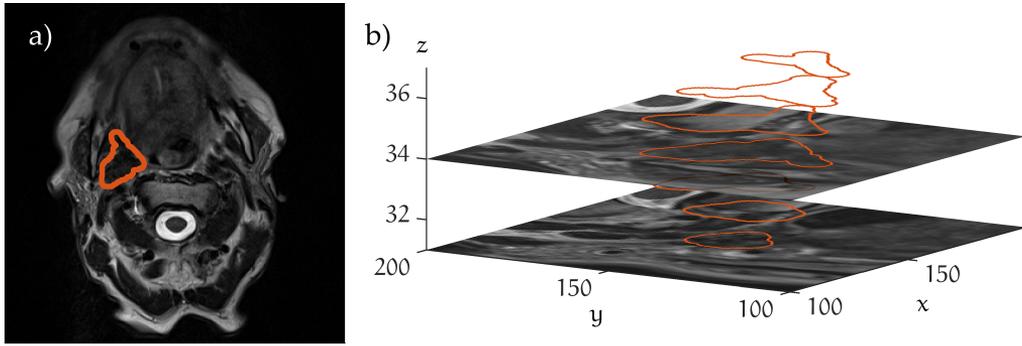


Figure 5.2: Tumor segmentations as polygons. a) shows a segmentation as drawn by an expert on a single image slice of a T_2 -weighted image. b) shows the same contours as stacked polygons overlaid on two example slices. As can be seen, there is no connection between the polygons in axial direction. Axis ticks denote voxel indices.

contour appears on all other available image sets. In this study, with only a few exceptions, the regions were defined on the T_2 -weighted image set.

For further processing of the data, the contours were transferred to MATLAB. The contours were exported as RT-structure files, which are defined by the DICOM standard [DIC21]: Each RT-structure file contains the contours as sets of multiple points linked to a reference coordinate system. The points define planar polygons in the coordinate system of the image set in which they were drawn, as depicted in figure 5.2. Additionally, affine transformation matrices M are stored in the RT-structure file that connect two coordinate systems Σ and Σ' , with

$$M = \begin{pmatrix} X_x d_i & Y_x d_j & Z_x d_k & S_x \\ X_y d_i & Y_y d_j & Z_y d_k & S_y \\ X_z d_i & Y_z d_j & Z_z d_k & S_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.1)$$

and

$$\begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} = M \cdot \begin{pmatrix} i \\ j \\ k \\ 1 \end{pmatrix}. \quad (5.2)$$

Here, (i, j, k) are pixel indices in the original image (Σ) and $P_{(x,y,z)}$ are the positions in the output coordinate system Σ' . (X_x, X_y, X_z) and (Y_x, Y_y, Y_z) are the row and column direction cosines of the image orientation, as stated in the DICOM standard [DIC21]. (Z_x, Z_y, Z_z) is then given by the cross product of X and Y . (d_i, d_j, d_k) are the voxel resolutions (including spacing in between

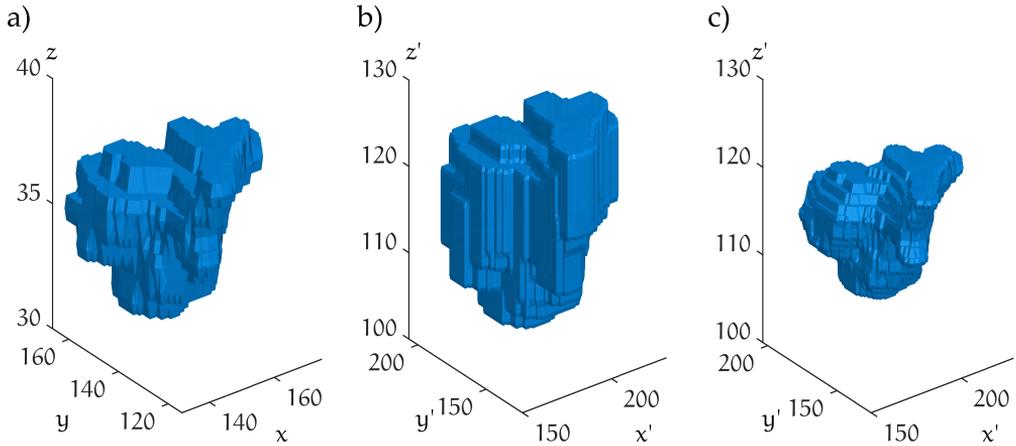


Figure 5.3: 3D rendered tumor segmentation masks. a) Original contour as defined by the radiotherapist, given in the T_2 reference space ($0.68 \times 0.68 \times 4.8 \text{ mm}^3$). It shows the same contour as fig. 5.2. b) Result of standard warping of the logical mask from a) to a resolution of $0.45 \times 0.45 \times 2.0 \text{ mm}^3$. c) Warping with linear interpolation in axial direction to the same reference space as in b).

the voxels) for each dimension, while (S_x, S_y, S_z) define translations. With the transformation matrix, any point in one coordinate system can be transferred to another coordinate system by application of equation (5.2).

However, the polygon points need to be transformed into a logical mask on a matrix grid. While, with a set of polygons, it is easy to create a 3-dimensional logical mask for the reference image set, the problem can become complex (i.e. inaccurate) if the mask has to be transferred to another coordinate system.

This interpolation problem can be split into two parts. First, the number of slices is increased in through-plane direction, which requires a careful interpolation strategy. Second, rotation of the new coordinate system with respect to the reference system leads to oblique polygons in the new coordinate system, preventing slice-based treatment of the problem. In an initial attempt, the logical region mask was warped to the new coordinate system using nearest neighbor interpolation. While this works well for voxels within the same plane, it shows staircasing artifacts in axial direction. Figure 5.3 a) and b) show this behavior with the original and warped versions of a logical mask.

To overcome this artifact, the interpolation between adjacent polygons was treated as a linear interpolation between two distance maps. A distance map $D(\vec{x})$ assigns the value of the closest distance to the edge of the polygon to each

pixel within the plane of the polygon (2D problem). Pixels inside the polygon are assigned negative values and those outside are positive, such that

$$D(\vec{x}) = \begin{cases} \min_{\vec{p}_i \in P} (\|\vec{x} - \vec{p}_i\|) & , \forall \vec{x} \text{ outside polygon } P \\ -\min_{\vec{p}_i \in P} (\|\vec{x} - \vec{p}_i\|) & , \forall \vec{x} \text{ inside polygon } P \end{cases} . \quad (5.3)$$

For any point in between the two polygons, a linear combination of the corresponding distance maps yields a new map:

$$D'(\vec{x}) = \frac{\partial z - \partial z_1}{\partial z} D_1(\vec{x}) + \frac{\partial z - \partial z_2}{\partial z} D_2(\vec{x}) , \quad (5.4)$$

where ∂z is the absolute distance between two slices and ∂z_1 and ∂z_2 are the distances of the point \vec{x} to the slice directly above and below. The interpolated logical mask is then a thresholded version of this new distance map $D'(\vec{x})$.

In the case of rotated coordinate systems, the described interpolation method can not be applied directly, since the polygons are oblique with respect to the new image planes. However, the same general strategy can be applied by treating each voxel in the new frame of reference separately: For this, each voxel is projected into the (oblique) planes of the polygons that lie directly above and below the voxel. Next, for each voxel, equations (5.3)-(5.4) can be applied, using the distances of the projected locations to their respective polygons. Figure 5.3 c) shows the result of the method.

In comparison, the linear axial interpolation method results in a smoother and much more confined volume, as can be seen in Figure 5.3 c). The nearest neighbor interpolation is unable to achieve results below the original voxel resolution, which explains the staircasing artifact.

5.1.3 MEASUREMENT SETUP

For all patient measurements, a clinical 3T whole-body MRI system (Tim Trio, Siemens, Erlangen, Germany) was used (figure 5.4). Each patient was provided with an individually fitted stereotactic radiation therapy mask that covered the complete head and part of the neck (figure 5.5). The mask is used to enable robust reproducible patient positioning as well as patient immobilization. Thus, the shared imaging information between CT, PET and MRI is nearly free from non-rigid deformations in the neck area. For the MR system, a special fitted headrest compatible with the masks is available, which allows seamless attachment of the mask directly to the patient table.

However, this setup prevents the use of a specialized head and neck receive coil. Therefore, flexible 4 and 6 (4-channel flex coil and Body Matrix Coil, Siemens Healthineers, Erlangen, Germany) channel receive coils are wrapped



Figure 5.4: 3 T whole-body MRI system (Tim Trio, Siemens, Erlangen, Germany) used in the FMISO trial.



Figure 5.5: Thermoplastic fixation mask as used in the FMISO trial.



Figure 5.6: Patient in the FMISO setup. Flexible signal receive coils are placed on the head and the chest of the patient, because a dedicated head and neck coil is incompatible with the thermoplastic fixation mask.

SEQUENCE	TE [ms]	TR [ms]	RESOLUTION [mm ³]	COMMENTS / OTHER
T ₁ FSE	11	504	0.7×0.7×4	
T ₂ FSE	100	5000	0.7×0.7×4	
Multiecho GRE	5-33	600	1.1×1.1×3	nEchoes=12, reconstructed map: T ₂ *
ss DW-EPI	71	3500	2×2×3	b={50, 400, 800} s/mm ² reconstructed map: ADC
rs DW-EPI	51	2510	2×2×3	b={50, 400, 800} s/mm ² reconstructed map: ADC
DCE	1.56	4.65	1.4×1.4×3	nTimepoints=36, reconstructed maps: K ^{trans} , k _{ep} , v _e
T ₁ VIBE Dixon	2.45	8.67	0.45×0.45×2	Post contrast. Water image used.

Table 5.1: Sequence parameters of the FMISO MRI protocol.

around the anterior part of the head and neck area and are used in combination with additional spine array coils, as seen in figure 5.6.

The tight wrapping and immobilization due to the mask can lead to increased patient heating and result in additional discomfort and stress which, in extreme cases, lets the patient stop the measurement completely. Hence, the measurement protocol has to be kept as short as possible to increase patient compliance and reduce additional motion, such as tongue motion and swallowing.

5.1.4 MEASUREMENT PROTOCOL

The MR protocol of the FMISO trial included 7 different sequences: anatomical T₁- and T₂-weighted fast spin echo (FSE), multiecho gradientecho (GRE), diffusion-weighted (DW) echo planar imaging (EPI), dynamic contrast enhanced (DCE) T₁-weighted VIBE and a post contrast T₁-weighted VIBE Dixon measurement. For the diffusion weighted imaging, two alternative sequences could be used: either conventional single shot (ss) or readout segmented (rs) EPI. Details on the most important sequence parameters are listed in table 5.1.

The T₁ and T₂ weighted FSE sequences are routinely used in many clinical MRI protocols, as they deliver high quality, high resolution anatomical reference images for diagnostic purposes, especially in tumor imaging.

For diffusion imaging, conventional ssEPI has been integrated in the measurement protocol as a gold standard. An additional rsEPI sequence was added to evaluate improvements in image quality, especially with regard to image artifacts such as distortion and Nyquist-ghosts. The differences of conventional ssEPI vs. rsEPI are discussed in section 5.2.

T_2^* measurements were performed using a multiecho GRE sequence to evaluate T_2^* as a possible substitute marker for hypoxic areas in tumors. The influence of blood oxygenation on T_2^* is well known from functional brain imaging, where it is called the BOLD (blood oxygen level dependent) effect [Oga90]. This effect relies on magnetic properties oxygenated and deoxygenated blood, i. e. due to diamagnetic oxyhemoglobin becoming paramagnetic deoxyhemoglobin.

5.2 SINGLE SHOT EPI VS. READOUT SEGMENTED EPI

The FMISO protocol (section 5.1.4) also includes two separate sequences for ADC quantification: conventional diffusion weighted ssEPI and rsEPI. ADC values from both sequences are compared in GTV-T and GTV-Ln in a region based approach. Therefore, ground truth volumes as defined in section 5.1.2 are used in combination with the original co-registrations as defined in the radiation therapy planning system. No additional region-based co-registration is employed.

Additionally, image quality with respect to apparent image artifacts is analyzed. The image artifacts under investigation are Nyquist ghosts and geometric distortions, that can lead to signal cancellation or signal voids due to ill-defined Fourier relationships (i. e. overlapping resonance frequencies for different spatial positions). Figure 5.7 shows examples of patients with good, medium and bad diagnostic quality, including the 4-fold Nyquist ghost and distortion artifacts in both rs- and ssEPI images.

The conventional ssEPI and rsEPI sequences were performed using a partial Fourier factor of 6/8 and an acceleration R factor of 2 (GRAPPA) in both cases. The resulting effective pixel bandwidths (see equation (3.14)) in phase encoding direction are $BW_{PE} = 19.5/48.8$ Hz. From equation (3.15), an inverse dependence between pixel bandwidth and geometric distortion is expected. To confirm that the geometric distortion is caused solely by B_0 inhomogeneity, a B_0 field map was measured in a volunteer experiment.

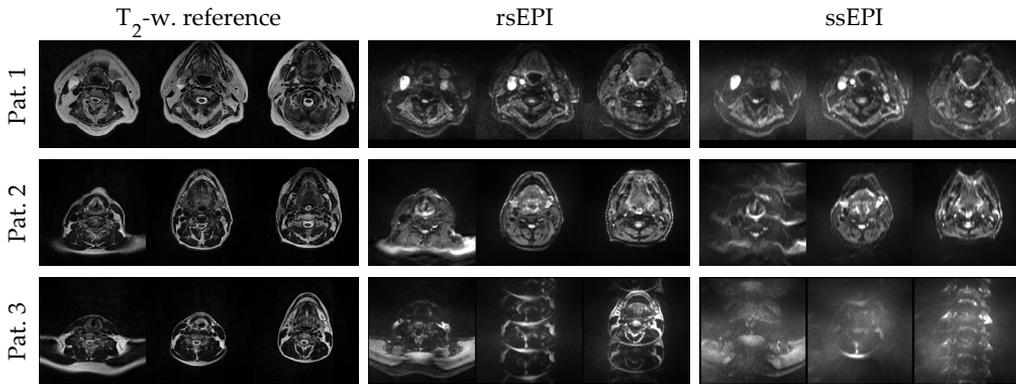


Figure 5.7: Examples for different qualities of rsEPI and ssEPI with respect to the Nyquist ghost artifact. The top row shows a grade 2 (good) example, the central row is rated 2 for rsEPI and 1 for ssEPI, and the bottom row is rated 0 (bad quality) for both. The three slices represent the most inferior, central and superior slice of the acquisition.

5.3 B_0 FIELD HOMOGENIZATION

As discussed in the previous section, imaging of the head and neck region is prone to artifacts from B_0 inhomogeneities. Large scale B_0 inhomogeneity can be prevented by B_0 shimming, which typically allows to compensate field distortions up to second order in x and y direction and up to fifth order in z -direction (along B_0) [Web16, p.166-207]. The specific geometry of the head and neck area, consisting of the head, smaller, approximately cylindrical neck and wide shoulders with air-filled lungs, imposes field distortions that typically cannot be compensated sufficiently by the shim gradients. This is especially critical in sequences with long TE and low bandwidth, such as diffusion weighted EPI.

To overcome the limited shim capabilities, liquid filled water bags were placed close to the region of interest in an attempt to simplify the complex geometry, similar as in [Oud16]. In volunteer measurements, a single U-shaped pillow was placed around the anterior part of the neck and resting on both sides on the patient table. The pillow was filled with pineapple juice, which closely matches tissue susceptibilities due to its high manganese content. B_0 field maps with and without the pillow were acquired using a 2-point GRE sequence (Siemens B_0 field mapping) [SG91] to quantify improvements in B_0 homogeneity.

Following the volunteer experiments, the patient measurement setup was adapted to account for additional limitations: Therefore, two commercially available hot water bags (figure 5.8) were placed on both sides of the head. The bags were filled with tap water doped with a gadolinium based contrast agent



Figure 5.8: Hot water bag used to reduce the geometric complexity of the head and neck area. Two water bags were filled with water doped with contrast agent and placed to both sides of the patients head.

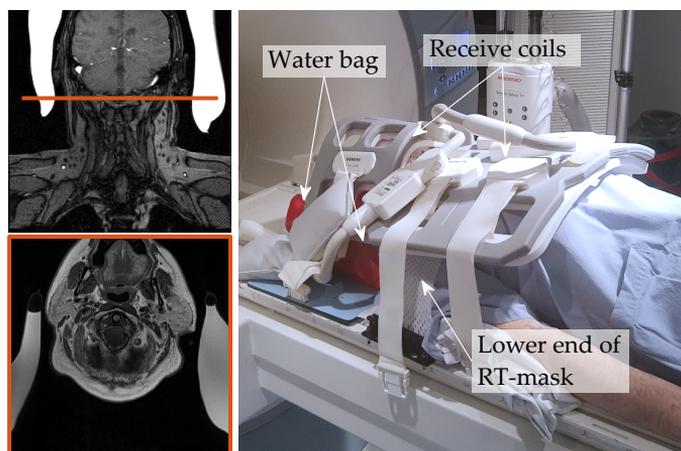


Figure 5.9: Patient setup in combination with anatomical MRI. The water bags are positioned as close as possible to the patient, but the thermoplastic fixation mask (white mesh seen in the photograph) prevents direct contact.

(ProHance 0.5 M, Bracco Imaging Deutschland GmbH) to reduce T_1 (20 ± 1 ms) and T_2 (16 ± 1 ms) to make them invisible in all sequences with a TE of more than 50 ms. Gadolinium doped water was preferred to the previously chosen pineapple juice due to its considerably longer storage times. This setup created a similar geometry as the U-shaped pillow did, without placing any weight on the anterior part of the patients neck, which could not be tolerated due to the proximity to the diseased area. The complete patient setup including an axial and a coronal MR image slice in the head and neck region is shown in figure 5.9.

An additional volunteer measurement in the patient setup, including the thermoplastic fixation mask, was performed to quantify the expected improvements on B_0 homogeneity.

Out of 29 patients in the FMSIO study, 9 patients were measured using the improved setup with the water bags. An analysis comparing the amount of Nyquist artifact within ssEPI and rsEPI before and after adding the water bags was then performed. The analysis uses a 3-point scale, consisting of grade 2: good image quality without Nyquist artifact, grade 1: acceptable image quality with some artifact in the outer, non-tumor bearing, slices and grade 0: diagnostically unusable due to corruption of Nyquist artifact. Figure 5.7 shows examples for the grading of the images. Because single patients are imaged multiple times during their therapy, a total of 20/22 measurements using the new setup were performed with ssEPI/rsEPI respectively, which were compared against 47/54 measurements with the original setup.

5.4 SEGMENTATION ALGORITHM

This section describes the details of the algorithm, especially the CNN architecture, that was used for tumor segmentation. Two CNN experiments using the CNN which are described in detail in this thesis: (1) the analysis of influence of distortion correction on the segmentation algorithm (section 5.5) and (2) the analysis of input channel contribution to the segmentation (section 5.6). Between these two experiments, additional patients were included in the study, and some of the methods of the CNN itself evolved. Therefore, this section will start with common methods for both experiments in subsection 5.4.1, while details on the specific parameters used for each experiment are found in sections 5.5 and 5.6.

5.4.1 CNN ARCHITECTURE

For all subsequent experiments, the general CNN architecture of *DeepMedic* as defined by [Kam17] has been chosen for segmentation. The basic building blocks of the network consist of a 3D-convolution layer, a batch-normalization layer, the activation layer (i. e., a leaky ReLU-layer) and, finally, a dropout layer.

The *DeepMedic* architecture does not utilize padding after each convolution layer, resulting in constantly decreasing feature map dimensions (see equation (4.7)). Additionally, smaller image patches cropped from the full image are used as input to the network. The decrease of feature map size and the use of image patches both lead to a receptive field which is much smaller than the original input images. Therefore, to segment the whole MRI volume, results from several patches have to be combined in a checkerboard fashion.

The architecture has a unique way of combining information of different scales of resolution, distinguishing it from other classical encoder networks: in a

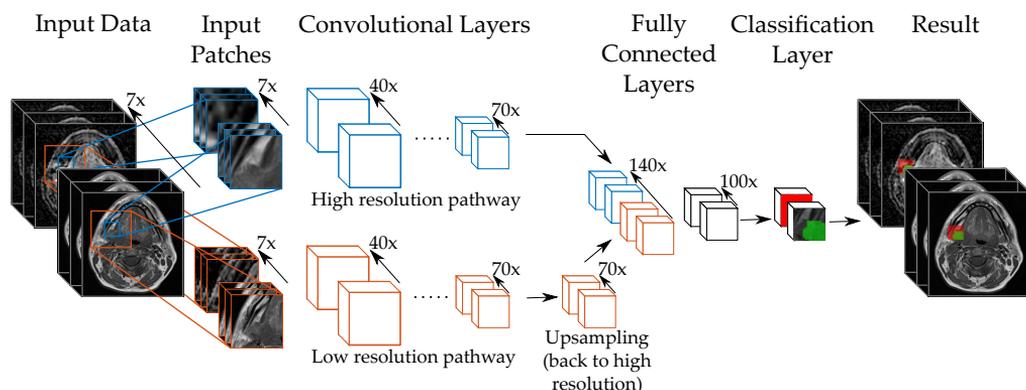


Figure 5.10: CNN Architecture of the *DeepMedic* implementation. The input is split into a high- and a low-resolution pathway, which are recombined before classification. Image adapted from [Bie18b, fig. 1], concept from [Kam17, fig. 5].

first step, the input patch is chosen, together with a larger patch centered around the input patch. The larger patch is then downsampled by a user specified factor (e. g. $2 \times 2 \times 1$ for x , y and z direction). The high-resolution, small-FOV patch (figure 5.10 blue), and the low-resolution, large-FOV patch (figure 5.10 orange) then enter separate pathways with a chosen number of convolution layers of the basic building blocks. Next, the low-resolution pathway is interpolated to the original resolution and concatenated to the output of the high-resolution pathway. In doing so, the low-resolution patches are cropped such that all pixel locations in the high- and low-resolution feature maps coincide. Then, information of the concatenated feature maps are merged by one or more fully connected layers, which are realized as convolutions with kernel sizes of 1^3 . In a final step a conventional classification using a softmax layer is applied. Figure 5.10 shows the full schematics of the architecture. The use of multiple pathways with different resolutions allows large scale contextual information to be incorporated in the segmentation process, without the need for overly excessive computer memory requirements.

The input to the network are multiple 3D MRI volumes (input channels). Each channel is a 3D image data set from a separate sequence or a parameter map. For the two experiments, different input channels were used: experiment 1 (section 5.5) used 5 and experiment 2 (section 5.6) used 7 unique channels. The network output can also be multiple channels, depending on the training data. If trained for *background*, *GTV-T* and *GTV-Ln*, there will be 3 output channels. The final softmax layer converts these pseudo-probability maps from the last convolution layer into binary segmentation maps.

To refine the results, the conditional random field (CRF) as proposed by [Kam17] and [KK11; KK12] was applied to the CNNs segmentation results. The application of the CRF decreases the amount of outliers, i. e. small false positive segmentations far from the lesion.

5.4.2 DATA PRE-PROCESSING

In section 5.1.2 the specific pre-processing of the ground truth segmentations was presented. In this section, general pre-processing of all available MR image data is described. As listed in table 5.1, from up to 7 different pulse sequences up to 9 different image contrasts and tissue parameter maps are extracted. For the two experiments, different subsets of this data were used as input to the segmentation algorithm.

The subject of data pre-processing can be split into two separate topics: data registration and interpolation on the one hand, and data normalization on the other.

Data registration is necessary if patient movement occurs in between scans of the same imaging session. Due to the fixation of the patient, the FMISO database contained only minor bulk registration errors. These were either corrected by an additional registration step in the whole MATLAB pre-processing chain (section 5.5.3), or during contouring using the IPlan software (section 5.6.2). Data interpolation deals with the different orientations and resolutions of the various MRI sequences. To process multiple MRI contrasts in a CNN, they need to be sampled on the same matrix grid. The DICOM information, which is stored with the image data, contains the orientation and location information of each image volume. Using this geometric information, data is gridded to the chosen reference frame, typically using cubic spline interpolation, with common matrix size, orientation and resolution.

Data normalization is a necessary pre-processing step in most segmentation tasks using CNNs. Contrast-weighted (e. g., T_2 , T_1 , FSE, or T_1 VIBE Dixon) sequences yield images without quantitative information to their absolute magnitudes, and, therefore, the dynamic range of the data can be different across different images. To make them comparable across patients (and possibly different MRI systems, coil setups, etc.), normalization to a common mean and standard deviation is necessary. The proposed normalization implicitly assumes identical positioning of the FOV and comparable anatomy across various patients. In the case of the FMISO study, this assumption is valid due to the rigid patient setup and measurement protocol, in combination

with a homogeneous patient cohort. Most commonly found is the z-score normalization

$$S'(\vec{x}) = \frac{S(\vec{x}) - \mu}{\sigma} ,$$

which transforms an image $S(\vec{x})$ with mean μ and standard deviation σ to $S'(\vec{x})$ with unit standard deviation and zero mean. For parameter maps, the physical meaning with respect to the absolute values should be conserved, particularly across the patient cohort. Therefore, constant values μ' and σ' must be extracted from the total patient cohort, which are the same for all images for normalization. This procedure is described in more detail in section 5.6.2.

5.5 DISTORTION CORRECTION

In the first major CNN based experiment (CNN-DistCorr), the influence of distortion in the ADC maps on the segmentation performance of the CNN was investigated. As discussed in sections 5.2 and 5.3, and described in 3.2.3, diffusion-weighted imaging can suffer from severe geometrical distortion. Therefore, a method for retrospective distortion correction without the knowledge of the true B_0 field map has been developed and applied to all available ADC data. Then, CNNs were trained based on five MRI input contrasts. These CNNs were trained from scratch, either with distortion corrected ADC maps or with original (distorted) ADC maps. Finally, segmentation results from the CNNs were compared to assess the influence of the distortion correction.

5.5.1 RETROSPECTIVE DISTORTION CORRECTION

The employed distortion correction algorithm [Bie18a; Bie19b] is based on the Lucas-Kanade method in a pyramidal layout [Bou00], which is, e. g., used in optical microscopy, 2-photon imaging or particle imaging [LK81; Vin14; PG17; GK09; AH10g]. Non-rigid motion is estimated between a T_2 -w. acquisition as reference and a potentially distorted diffusion-weighted image with lowest available b-value. A schematic of the algorithm is shown in figure 5.11. A rigid registration (translation only) is found by optimization of the mutual information metric between the distorted and the reference image [VI97]. In the next step, the registered image is broken into smaller patches. Unlike in figure 5.11, the image is broken into 9 (2D) or 27 (3D) patches, which are overlapping by half their side length in each dimension, to account for the continuous properties of naturally occurring field distributions. Each patch is then registered with its corresponding patch from the reference image, which

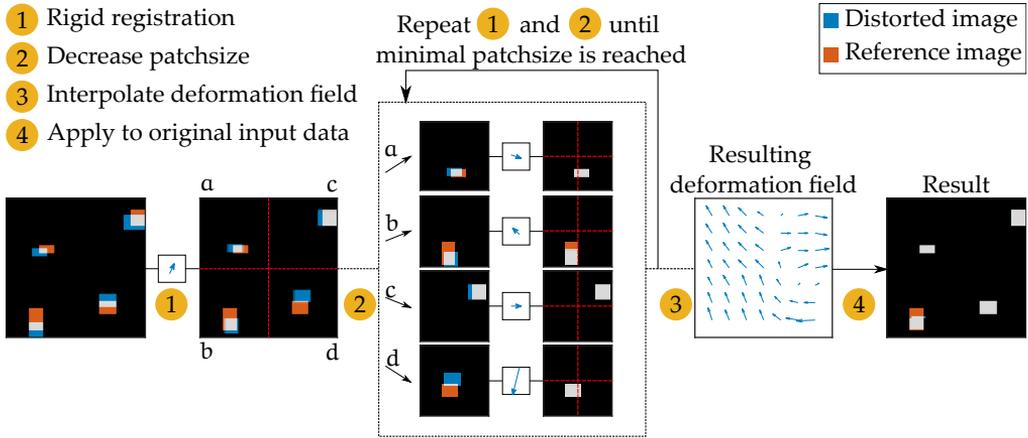


Figure 5.11: Schematic of the distortion correction algorithm. Given a distorted (blue) and a reference (orange) image, the algorithm employs mutual information based rigid registration on the full image. After registration, the image is divided into smaller, overlapping patches (not shown in the figure). The smaller patches are registered again, and are sub-divided subsequently. Registration results are passed on to the next level for patch selection. Each pyramid level decreases the patch size by 2^n , where n is the image dimension (2 or 3). When a preset limit is reached, the resulting distortion field can be used to correct the deformed image. Figure adapted from [Bier8a, fig. 1]

leads to registration features with finer resolution than in the previous step. This procedure is repeated until a predefined minimal patch size is reached. The final distortion field is given by interpolation of the finest features to the original matrix size.

In this study, the minimal patch size was set to $10 \times 10 \times 10$ pixels. Tests on smaller patch sizes did not yield more accurate results, as the registration becomes unstable due to the small number of pixels per patch. The algorithm was implemented for 3D data in MATLAB. Since the distortions are only expected to occur along the phase-encoding direction, the algorithm was implemented such that the registration was limited to a single spatial dimension. This was achieved by ignoring translation in dimensions different from the selected dimension during steps 1 and 2 in the algorithm, thereby ensuring a smooth convergence of the registration. After the distortion field is estimated from the reference T_2 -w. image and the DW image it is applied to the corresponding ADC map. Since the amount of distortion should only depend on the echo train length, which is equal in all diffusion-weighted images with differing b -values from the same acquisition, the use of the distortion field can also be applied to the calculated ADC map. The lowest b -value image was chosen as a basis as it has the highest SNR and is most comparable in contrast to the T_2 -w. image.

To validate the algorithm performance and the assumption, that distortion in DW MRI is mainly caused by B_0 inhomogeneity, diffusion weighted and T_2 -w. images of a volunteer were acquired in the head and neck region with the same protocol as in the patient measurements. Additionally, a B_0 field map was acquired using a multiecho gradient echo sequence. From the DW and T_2 -w. images, a distortion map was calculated using the retrospective correction algorithm. The resulting distortion field was then compared against the measured field map.

5.5.2 AVAILABLE DATA

For this experiment, a total of 18 complete datasets were available with 5 input channels to the model: T_1 - and T_2 -FSE and K^{trans} -, T_2^* - and ADC- parameter-maps. From the available data, 12 out of 18 ADC maps were calculated from the rsEPI sequence, while for the rest only ssEPI data was available.

5.5.3 DATA PRE-PROCESSING

T_1 and T_2 weighted data were subject to a z-score normalization. The normalization was done per 3D volume, and mean and standard deviation were calculated only within a region mask. The mask was created by a simple thresholding of the image, setting all values below 10% of the maximum intensity to 0, and all others to 1. The threshold value was found experimentally by visual inspection of the results: when most of the empty volume (i. e., air) was masked and the anatomy was still completely present, the threshold value was accepted.

To preserve the physical properties of the parameter maps, no normalization was applied to K^{trans} , T_2^* and ADC.

All imaging data was interpolated to a common isotropic resolution of 1 mm^3 and additionally registered to the T_2 -w. image using the mutual information metric limited to similarity transformations (i. e., translation, rotation and scale). After registration, all image sets were checked visually for potential misregistrations. Additionally, only regions containing information from all 5 input images were used, i. e., proper masks were applied to the data. Therefore, for each of the 5 channels a specific mask was generated by thresholding as described before. Then, a combined mask was calculated as the intersection of all single masks. The resulting combined mask was then applied to each single channel.

The retrospective distortion correction algorithm was then applied to the ADC maps. As a result, two sets of input data for the subsequent CNN were available:

- the *uncorrected set*: T_2 , T_1 , K^{trans} , T_2^* and original ADC,
- the *corrected set*: T_2 , T_1 , K^{trans} , T_2^* and distortion corrected ADC.

5.5.4 CNN ARCHITECTURE

The general layout of the CNN has been discussed in section 5.4.1. For this experiment, the original Python model *DeepMedic Vo.6* [Kam17] was used and hyperparameters were adapted [Bie19b].

The network output was either GTV-T or non-GTV-T (i.e. background). Thus, lymph node metastases were counted as background in this task. One subsampled resolution pathway was chosen with a factor of $3 \times 3 \times 3$ lower spatial resolution in x , y and z direction in addition to the high-resolution pathway (1 mm³ isotropic). Each pathway included 8 hidden layers, each with {30, 30, 40, 40, 50, 50, 70, 70} channels, followed by two fully connected layers with 100 channels reconnecting the two pathways. Following [Kam17], convolution kernel sizes in the two pathways were chosen as 3^3 . During training, the input data was segmented into patches of size $20 \times 20 \times 20$ and batches of 10 patches were processed before updating network weights. For validation and inference, combinations of larger patches and/or batch sizes were chosen to leverage the computing power, but they do not influence the training outcome. The last fully connected layer and the classification layer used a dropout rate of 50%. To account for the inherent class imbalance, patches centered on GTV-T and background were chosen with a relative frequency of 0.5/0.5 and a random mirroring of the input data in each dimension was performed for data augmentation. The initial learning rate was set to 0.001 and was decreased by a factor of 2 on epochs {10, 14, 17, 20, 23, 25, 29, 31 and 34}. In total, the network was trained for 35 epochs. Figure 5.12 shows the training progress for a network trained on the described parameters. Detailed configuration files are found in the appendix 11.1 and 11.2.

Parameter optimization was performed by manually exploring the parameter space. The largest influence on training outcome was given by the depth and width of the network architecture, i.e. the number of hidden layers and their respective number of channels. Deeper networks were expected to generate more precise segmentations, but with more than 8 hidden layers the results became increasingly noisy. The number of channels per layer and the batch size/patch size were then optimized together, as both are limited

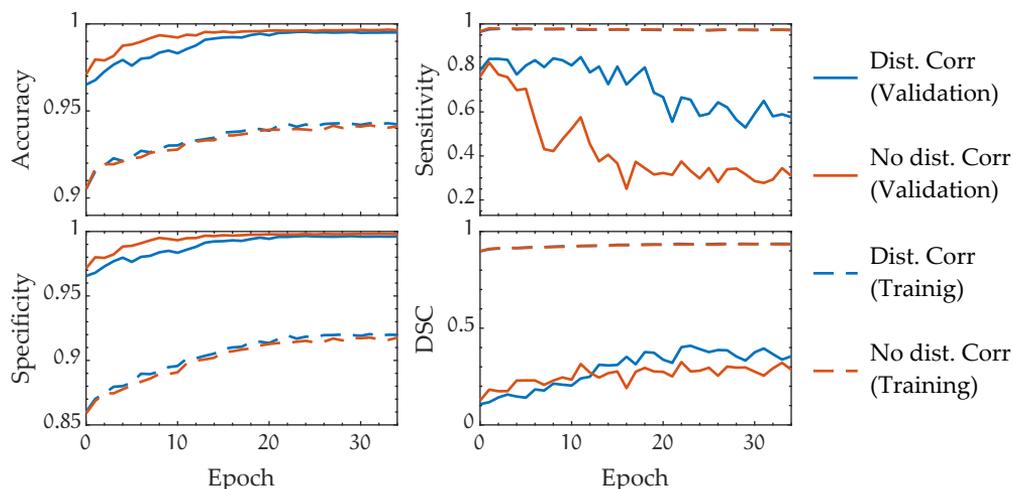


Figure 5.12: Training progress for the network. Figure adapted from [Bie19b, fig. 2].

by GPU memory. While the number of channels per layer is expected to increase the power of generalization, the batch size had to be chosen sufficiently large to catch a statistically representative part of the tumor and background distributions in each optimization step.

5.5.5 EXPERIMENT DESIGN

The experiment was designed to identify whether distortion correction of ADC maps improves tumor segmentation of a CNN. From the training curves (figure 5.12) in the CNN testing phase, the need for a larger database to discriminate between subtle changes became apparent. Thus, with 18 valid patient data sets in total, a 14-fold leave-one-out cross validation strategy was employed: the complete dataset was split into subsets consisting of a training set (13 patients), a validation set (4 patients) and a test set (1 patient). 14 different permutations of these subsets were generated, each permutation with a unique patient in the test set. For the corrected and uncorrected dataset, as described in section 5.5.3, the same permutations were chosen to keep them comparable.

Then, for each of the 14 permutations, a segmentation network was trained from scratch for uncorrected and corrected data. After complete training of the networks, the corresponding test sets were segmented and results between the corrected and uncorrected datasets were compared. Statistical analysis, including a paired students t-test was employed to reveal statistically significant differences between the two sets.

As a performance marker, the DSC was chosen predominantly, because of the robustness against the relative size of the target region with regard to the total image size.

5.5.6 COMPUTATIONAL RESOURCES

The high computational load for CNN training demanded a local server infrastructure with up to 4 NVIDIA Tesla C2075. Training times for 35 epochs on the given hardware were between 50–75 hours per GPU, depending on the total computational load on the server.

5.6 CHANNEL-WISE IMPORTANCE ANALYSIS

The second major CNN based experiment (CNN-Info) deals with the relative importance of each input channel towards the overall segmentation performance of a CNN [Bie20b]. Since the first experiment CNN-DistCorr, (section 5.5) the patient database as well as the CNN methods were expanded. Similarly to the previous section, the available data will be described first, followed by the CNN methods. Importantly, transitioning from the first to the second experiment, the whole CNN structure was ported to the MATLAB framework. This allowed a more direct access to the methods used in the CNN and changes in the design to be made. The porting and implementation of *DeepMedic* into MATLAB was accompanied by a collaboration project with MathWorks, giving the option to access neural network methods at the systems very basic implementation.

5.6.1 AVAILABLE DATA

Data for this experiment was taken from the same study as in section 5.5. Because of the prospective nature of the FMISO study, 6 additional patients were included totaling to 24 MRI patient data sets. With each patient receiving 3 MRI exams in the course of the therapy, 72 independent MRI were performed. From this, about half of the data had to be excluded due to artifacts (motion, B_0 inhomogeneity) or patient compliance (interruption of MR measurements due to discomfort). As a result, 36 complete MRIs, taken from 18 patients, were available [Bie20b].

For the CNN analysis, 7 different MRI input channels were chosen: T_1 and T_2 FSE, post contrast T_1 VIBE Dixon (water image) and T_2^* , ADC, K^{trans} and v_e parameter maps (see table 5.1). These input channels represent the complete MRI protocol. The perfusion parameter map k_{ep} was discarded because it does

not carry additional information according to the Tofts model [Tof97; Tof99], if K^{trans} and v_e are available.

5.6.2 DATA PRE-PROCESSING

For the ground truth segmentations data were imported into a radiation therapy planning software (see section 5.1.2). There, a registration of MR images based on the mutual information metric was applied, if necessary. The information generated by the registration in the planning software was subsequently exported alongside the ground truth segmentations. All data were warped to a new, common frame of reference. The choice for such a common frame of reference had to fulfill several requirements: First, the voxel resolution should be in the range of the data which will be used in the segmentation algorithm to avoid loss of information by excessive downsampling. Second, the orientation of the common frame of reference should be the same as at least one of the acquired image data sets to reduce the need for interpolation and associated artifacts (preferably, the least amount of images should undergo rotational interpolation). Finally, the common frame of reference should only include a field of view which is covered by all of the underlying images, to avoid region voids without any information from certain images. Unlike the ground truth volumes, the images are then interpolated using cubic interpolation, as integrated into the MATLAB software package. The T_2 FSE sequence was chosen as the reference image orientation, because it was used for manual ground truth contouring and the MRI protocol FOV was referenced to it.

The target resolution was chosen as the highest measured resolution in x,y and z directions, i. e. $0.45 \times 0.45 \times 2 \text{ mm}^3$.

Data normalization improves the convergence behavior of CNN training. The activation functions used in this experiment contain a nonlinearity at $x = 0$ (leaky ReLU). Using the He-initializer for weight initialization [He15], a normalization to the approximate range of $[-1, 1]$ increases learning speed by setting values around the point of largest gradients. Therefore, contrast weighted images as well as parametric images (i. e. perfusion K^{trans} and v_e , ADC and T_2^*) were normalized. Due to the physical meaning of the parameter maps, two different normalization strategies were applied. As mentioned above (section 5.4.2), all contrast weighted images were subject to a z-score normalization. However, during optimization of the network parameters, modified values of 0.25 each have proven beneficial for both, mean and standard deviation.

Normalization of parametric images was based on a histogram normalization to preserve the physical absolute values of each data point (voxel) [Jac19].

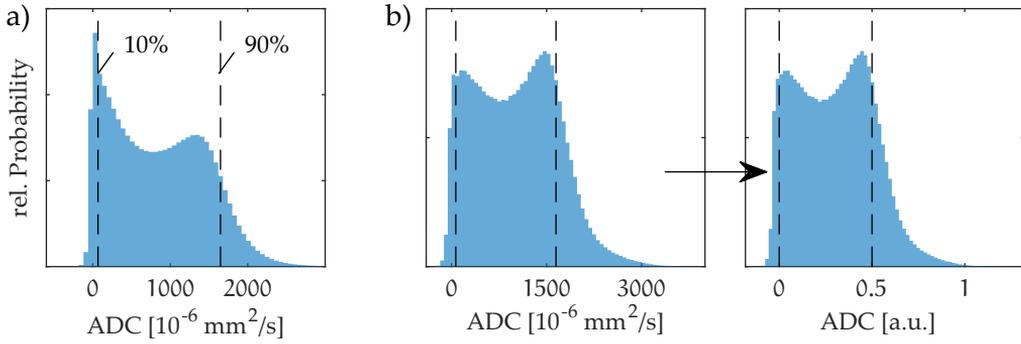


Figure 5.13: Histogram based normalization of ADC values in a) all available images and b) for one selected 3D image. The dashed lines represent the 10% and 90% quantiles with respect to the complete set of images. b) shows the histogram of one sample image before and after normalization. The data shown in the histograms exclude background, which has been removed using a region mask.

Therefore, all available data across all patients for a specific parameter (e. g. ADC) were collected in a single histogram, as shown in figure 5.13 a). Then, normalization parameters μ' and σ' were chosen such that data between the 10% (q_{10}) and 90% (q_{90}) quantiles fell into the range of $[0, 0.5]$:

$$\begin{aligned}\mu' &= (q_{10} + q_{90})/2 \\ \sigma' &= q_{90} - q_{10} \\ \text{ADC}_{\text{norm}} &= \frac{\text{ADC} - \mu'}{2\sigma'} + 0.25\end{aligned}$$

Note that after normalization the 10% and 90% quantiles are not equal to 0 and 0.5, respectively, for each individual image. This relation only holds for the mean of all images.

Normalization was performed after the 3D-volumes which were warped into the common frame of reference.

5.6.3 CNN ARCHITECTURE

For this experiment, the *DeepMedic* architecture was ported to MATLAB and some features were added.

A network architecture with 2 separate pathways was chosen. Due to the anisotropic resolution of the input images, the subsampling factor of the low-resolution pathway was $3 \times 3 \times 1$ in x , y and z direction. As input to the network, a single large patch which covers both, high and low resolution, is chosen. This input is then split into the two pathways, and a cropping layer makes sure that only the smaller patch for the high-resolution pathway is processed. Next, the

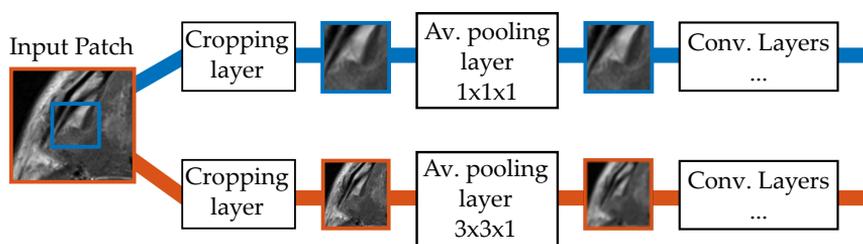


Figure 5.14: Schematic of the initial patch processing in the two-path network. The input patch is split into a smaller patch which is processed in the convolutional layers directly, and a larger patch which is subject to average pooling, first. Technically, the smaller patch is also average-pooled with a kernel of $1 \times 1 \times 1$, which is an identity transformation. However, setting up the network this way allows for more flexibility, e. g. if in some cases average pooling for the small patch is required, too.

downsampling was performed in an average-pooling layer with a pooling size and stride that matches the downsampling factor. For a factor of $1 \times 1 \times 1$, no pooling is applied. In this setup, a single input to the network serves for both the low- and high-resolution pathway. A schematic of the input processing is shown in figure 5.14.

Each pathway contained 10 convolution layers. In this setup with an anisotropic input resolution, the kernel sizes of the layers were chosen such that the receptive field would cover a physical area of approximately equal side lengths. Therefore, kernel sizes were $3 \times 3 \times 3$ for layers {1, 5 and 9} and were $3 \times 3 \times 1$ for layers {2, 3, 4, 6, 7, 8 and 10}. Applying equation 4.7, the resulting receptive fields for high- and low-resolution pathways are $21 \times 21 \times 7$ pixels, which equals to $9.45 \times 9.45 \times 14 \text{ mm}^3$ and $28.35 \times 28.35 \times 14 \text{ mm}^3$ respectively. Upon upsampling of the low-resolution pathway and concatenation of high and low resolution, another 3 convolution layers with kernel sizes $3 \times 3 \times 1$, $1 \times 1 \times 1$ and $1 \times 1 \times 1$ were added. The number of feature maps was 104 for each convolution layer in the separate pathways and 150 for the connected pathways. In this experiment, each layer consisted of building blocks with the following operations: first, a convolution layer with kernels and number feature maps as described above. Next, a learnable batch normalization layer [IS15], a leaky ReLU layer with a scale of 0.01 and finally, a dropout layer with a 20% dropout chance (see figure 5.15). For convolutions, no zero padding was applied and all strides were set to 1.

In section 5.5.4, the upsampling of the low-resolution pathway was realized as a simple nearest neighbor interpolation. In this implementation, the up-sampling operation was implemented as a learnable unpooling (or transposed convolution) layer. The layer was applied with a kernel size of $3 \times 3 \times 1$, matching the downsampling factor for the low-resolution pathway.

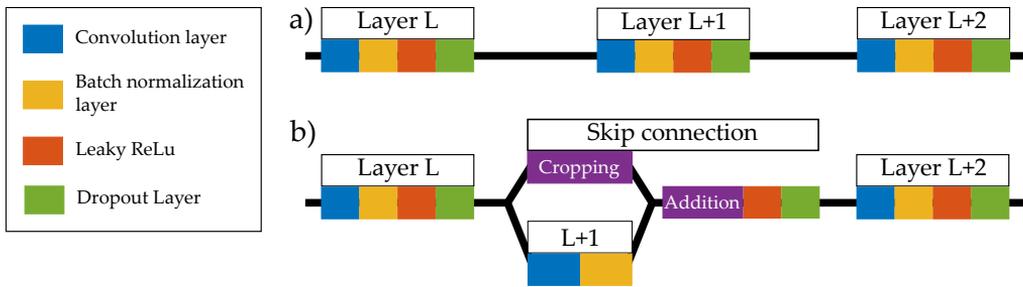


Figure 5.15: Schematic of the basic network structure. a) shows 3 connected hidden layers in succession. b) shows the implementation of a skip connection between layers L and L+1.

In the first experiment CNN-DistCorr, increasing the number of hidden layers eventually decreased the segmentation performance. To overcome this instability, residual connections (or skip connections) [He15] were introduced: Layers with skip connections add the output of a layer A to the output to a successive layer B, before the nonlinearity operation of B (see figure 5.15). This connection can improve the learning success for large numbers of hidden layers when the function connecting one layer to the next is close to identity. If that happens, introducing the skip connection changes the function $f(x) = x$ from identity to $f(x) = 0$. This is generally easier (i. e., faster) to approximate using a gradient based optimization. Typically, the feature maps in the skip connection layer are zero-padded to retain the matrix size of the previous layer. Here, a cropping was implemented to mimic the decreasing feature map size which is present throughout the network. This way, memory can be saved, because no average pooling is applied as in the original *ResNet* [He15].

Finally, a softmax layer generated segmentation maps for classes background, GTV-T and GTV-Ln. The network was then trained using the generalized Dice loss function [Sud17], i. e. MATLABs *dicePixelClassificationLayer* adapted for 3D input.

For training, input patches of size $78 \times 78 \times 8$ px were used, which were immediately cropped to a size of $38 \times 38 \times 8$ px for the high resolution pathway (figure 5.14). Accordingly, output patches of size $18 \times 18 \times 2$ px were generated.

A random sampling strategy was implemented for the selection of the patches: First, a probability map for each category is calculated from the data and the ground truth labels. Therefore a simple threshold based mask is generated from the MRI data to exclude low-signal areas, such as air. This mask is applied to the ground truth label maps (GTV-T, GTV-Ln, BG), and a subsequent Gaussian image filter with standard deviations of $\{10, 10, 0.01\}$ for x, y and z directions is applied. These maps are then vectorized (i. e. the number of dimensions

is compressed to one) and the cumulative sum across the vector is calculated and normalized to 1 (cumulative probability density). Next, a random category is chosen with a probability of $1/3$ each. Now, the central pixel of the image patch is chosen by calculating the difference of the cumulative probability density and a random number between 0 and 1. The pixel in question is given by the index with the smallest difference and the random number. This procedure has a significant advantage over a completely random choice of pixels: Picking the central pixel from the smoothed probability map allows the central pixel for a category to be outside of the region of that category itself. Therefore, small structures and fuzzy edges can be sampled more smoothly and a better training statistic is expected. At the same time, the probability for these choices is reduced, as it is weighted by the amplitude of the probability map. Additionally, areas containing only air are suppressed from the sampling.

Data augmentation was implemented by random flipping of the image in any dimension (probability of 0.5 for each dimension x , y or z) and a chance of $1/3$ for a random rotation in the x - y -plane. Rotations about the x - or y -axis were excluded due to the non-isotropic image resolution.

In summary, the network contains 2.78 million learnable parameters, 99.7% of which are the weights of the convolution layers. Other learnable parameters are the convolution layers bias (0.1%) and scale and offset parameters of the batch normalization layers (0.2%).

5.6.4 EXPERIMENT DESIGN

The goal of this experiment was to discriminate between MRI input channels that add a lot of information to the CNN for tumor segmentation and those that add very few or no information. Therefore, multiple networks using the same structure and the same input data were trained from scratch. Each network was trained on 6 out of 7 available input channels, cycling through all 7 possible configurations in a leave-one-out (LOO) fashion. Additionally, one network with all available input channels was trained for reference. As before (section 5.5.5), the size of the test data set was restricted to a single patient, thereby maximizing the training data set. This, too, was done employing a leave-one-out strategy, resulting in 288 completely trained networks for 8 different input channel configurations.

The results were compared using the DSC in terms of segmentation performance for GTV-T and GTV-Ln. Paired Students-t-tests were employed between the network trained on all available input channels and the LOO-networks, to assess statistical significance between them.

5.6.5 COMPUTATIONAL RESOURCES

The computational resources necessary for the experiment CNN-Info were not available within the local server structure. Therefore, Amazon Web Services (AWS) was used for the computation, which provided Linux-based host computers equipped with Tesla T4 GPUs (instance name: g4dn.xlarge). Other instances with more powerful GPUs could have been chosen, but the g4dn.xlarge was the most cost efficient variant. Using this, up to 6 machines were run in parallel, each training a different network.

For data protection, the transferred data was fully pseudonomized and formatted in the *nifti* file format, with all header information removed. The remote data storage was encrypted using AWS encryption technology with an 128 bit ssh key, with the private part only being available locally. The local data protection officer approved the procedure.

SIMULTANEOUS ADC AND T_2 MAPPING

Based on the results from chapter 5, diffusion ADC acquired with SE-EPI sequence suffered from severe artifacts, and T_2 is among the most important clinical contrasts. Therefore, a sequence for simultaneous ADC and T_2 mapping was developed, called SATM. The main goal was to design the sequence to overcome the image distortion and ghosting artifacts (5.3). Additionally, increased acquisition speed as compared to separate ADC and T_2 measurements is desired, with the additional benefit of intrinsic co-registration for ADC and T_2 maps.

6.1 SEQUENCE DESIGN

The proposed new pulse sequence SATM is based on a multiecho spin echo readout train with a radial k-space trajectory and interleaved diffusion sensitization blocks (6.1). A unique pair of TE and b-values is assigned to each echo in the echo train, since each refocusing pulse increases TE by a set amount ΔTE , and each diffusion block increases the b-value by Δb . The echoes are then sorted into their corresponding k-spaces, and the acquisition is repeated n_{sp} times. Additionally, the acquisition can be repeated with a different echospacing ΔTE to further increase number of unique b-TE-pairs.

The radial spokes in k-space are acquired in a golden angle fashion [Wino7], such that no angle is sampled more than once, and any set of consecutive spokes has a homogeneous azimuthal distribution. This method allows for sparse sampling of k-space, as undersampling artifacts are incoherent throughout the different b-TE images.

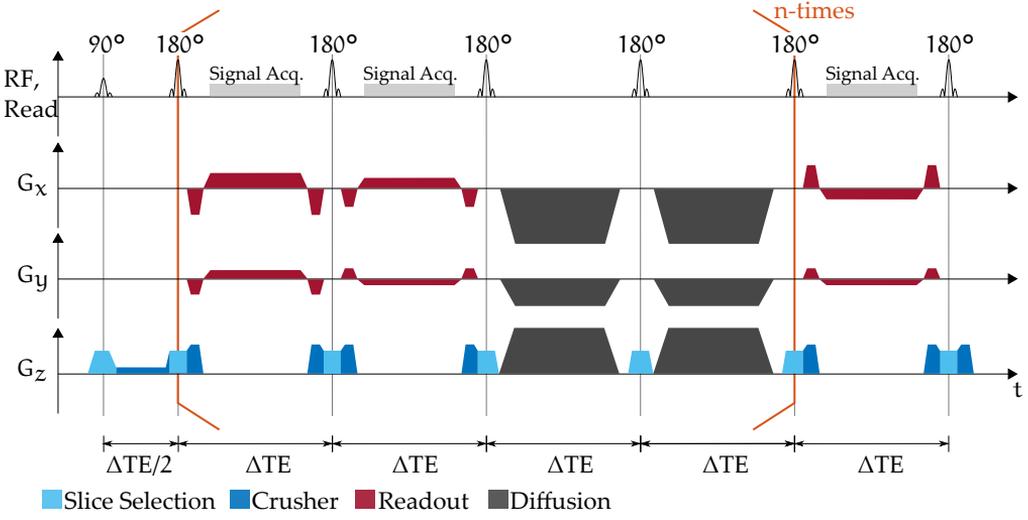


Figure 6.1: Sequence diagram of the SATM sequence. The building block, marked by the orange parenthesis, can be chosen with different numbers of readouts. Similarly, the diffusion block can be shifted to a different position, e. g. before the first readout. Figure adapted from [BLB21, fig. 1].

Crusher gradients are applied in slice selection direction to suppress FID signal contribution due to imperfect slice selection gradients [BE84]. The crusher gradient moment M_{cr} is given by

$$\begin{aligned} M_{cr,1} &= 0.5 \left(\frac{1}{2} M_{ss} \right) \\ M_{cr,2-n} &= 1.5 \left(\frac{1}{2} M_{ss} \right) \end{aligned} \quad , \quad (6.1)$$

where M_{ss} is the moment of the slice selection gradient, $M_{cr,1}$ is the crusher after the 90° excitation and $M_{cr,2-n}$ are all following crusher gradients.

Diffusion sensitization is achieved by interleaving Stejskal-Tanner diffusion blocks [ST65] with identical echo spacing ΔTE in the regular echo train. The interleaved diffusion blocks violate the CPMG conditions (see section 2.2), which can cause severe artifacts and signal voids in the resulting images [SH98]. To force all echoes to occur at the same time, in the center of each readout block, a soft CPMG condition is formulated: The integral of the gradients in each part of the diffusion block must be an integer multiple of the gradient integral of a regular readout block:

$$\int_{t_D}^{t_D + \Delta TE} G(\vec{x}, t) dt \equiv N \int_{t_R}^{t_R + \Delta TR} G(\vec{x}, t) dt \quad (6.2)$$

where t_D is the start time of a diffusion block and t_R is the time of the beginning of a readout block. N is an integer and $G(\vec{x}, t)$ are the gradients. Note that the gradient integrals of a readout block essentially consists only of the crusher gradients, as the readout gradients are completely rewound.

To obtain the largest possible diffusion b-value within a limited amount of time, all three spatial gradients are switched simultaneously. The resulting diffusion gradient is then given by a combination of the three physical gradient directions. Here, three new virtual gradient directions are constructed, which form a new basis Σ' . Three conditions for the new basis were enforced:

1. Σ' is orthogonal, and all basis vectors have the same length.
2. The z-component (through-slice component) is the same for all directions.
3. Given 1. and 2., the gradient amplitude is maximized.

Condition 1 ensures that a trace-weighted ADC can be calculated from three separate measurements, as it is done in conventional DW-EPI. Condition 2 guarantees equal conditions for through slice off-resonance effects, independent of the gradient direction. And last, condition 3 yields the largest possible b-value. Given this set of conditions, the new basis can be written as

$$\Sigma' = \begin{pmatrix} x' & y' & z' \\ 1 & -0.73 & -0.27 \\ 0.27 & 0.73 & -1 \\ 0.73 & 0.73 & 0.73 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} x & y & z \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (6.3)$$

where Σ is the canonical basis given by the physical gradient directions, and z is the slice selection direction for non-oblique slices.

Since eddy currents can lead to severe artifacts in multi-echo spin echo sequences, the rise time of the diffusion gradients was set to $t_{D,\text{rise}} = 1.5$ ms, resulting in a maximum slew rate of 46.4 T/m/s and a maximum gradient amplitude of 69.6 mT/m. The limit of 1.5 ms was found empirically as it showed minimal artifact but still allowed reasonable b-values. This slew rate is well below the nominal limits of the system with a maximum gradient of 80 mT/m and a slew rate of 200 T/m/s. Additionally, a variable dead time was introduced between the diffusion gradient and the next 180° pulse to eliminate rapidly decaying eddy currents.

To ensure good slice profiles, long excitation times of $t_{\text{Ex}} = 3\text{--}4$ ms were chosen for the experiments. Dedicated TSE pulse shapes were used to further improve profile sharpness [Pau91].

6.2 IMAGE AND PARAMETER MAP RECONSTRUCTION

In this section, the reconstruction of images and parameter maps (S_0 , T_2 , ADC) from the measured k-space data is described. Starting with the implementation of the NUFFT as the basis for all further calculations (section 6.2.1), two alternative reconstruction algorithms are presented (sections 6.2.2 and 6.2.3). Finally, sections 6.2.4–6.2.6 highlight the relevant data pre-processing steps for the reconstruction.

6.2.1 NON-UNIFORM FAST FOURIER TRANSFORM

Generally, in the reconstruction a transformation from a radially sampled k-space to a Cartesian image space is needed. Here, the transformation is realized using a gridding kernel, which interpolates k-space samples to a Cartesian grid given a set of measured k-space points at arbitrary but known locations. The Cartesian k-space can then be transformed to image space using the well-known Fast-Fourier transform. The implementation is based on code from [Fes] and includes an efficient interpolation mechanism which is based on Kaiser-Bessel interpolation kernels [Jac91; Fes07]. The whole interpolation is pre-calculated once and stored in a sparse matrix, which can be used on any k-space data on the same grid, thereby dramatically increasing calculation speed. Similarly, the inverse transformation from Cartesian to radial space is performed using the same toolbox. The Kaiser-Bessel window was 6×6 pixel and a 1.5 oversampling is used to increase the accuracy. An efficient Non-Uniform-Fast-Fourier-Transform (NUFFT) is needed because the reconstruction generally needs to perform hundreds, or even thousands, of independent transforms. In a typical example, with 10 iterations until convergence, a 20-channel head coil and an average of 3 iterations in the line search of the conjugate gradient algorithm, 1600 transformations between image and k-space are performed for a single slice. The typical calculation time for this example was around four minutes for a single slice with a matrix resolution of 128×128 .

6.2.2 MODEL-BASED RECONSTRUCTION

In this work a model-based reconstruction approach was implemented based on the works by [BUF09] who developed a model-based iterative reconstruction algorithm for T_2 mapping. Here, the measurement process is expressed in a model function \vec{F} , which includes coil sensitivities C_c , the NUFFT $A(\cdot)$ and a signal contribution which is defined by a physical model. The reconstruction then minimizes a loss function Φ , as described in section 2.3.2.

Here, the pixelwise signal model was extended to incorporate the attenuation due to diffusion and an additional, inconsistent phase Θ which can arise in diffusion imaging (section 3.2.3):

$$S(t) = S_0 e^{-b(t)D - t_E(t)R_2 + i\Theta(t)} \quad (6.4)$$

$R_2(x, y) = 1/T_2(x, y)$ is the relaxivity, $D(x, y)$ is the diffusion coefficient for each voxel, and $\Theta(x, y)$ the image phase for each b-TE-pair. The new loss function Φ then reads:

$$\Phi = \frac{1}{2} \sum_t \sum_c \|\vec{F}(\vec{x}, t, c) - \vec{y}(t, c)\|_2 + \lambda \|W(\vec{F}(\vec{x}, t, c))\|_1 \quad (6.5)$$

with

$$\vec{F}(\vec{x}, t, c) = A \left(C_c(\vec{x}) S_0(\vec{x}) e^{-b(t)D(\vec{x}) - t_E(t)R_2(\vec{x}) + i\Theta(\vec{x}, t)} \right) \quad (6.6)$$

Here, $\vec{y}(t, c)$ are the measured k-space data for echo times t and coil C_c and $W(\cdot)$ is the regularization transform with λ being a free parameter to balance the contribution of regularization. $A(\cdot)$ is the NUFFT operation and $C_c(\vec{x})$ are the complex coil sensitivity maps. For regularization, either total variation or wavelet transformation are used. Both regularization reliably methods suppress noise. Here, the wavelet regularization term was preferred to the TV term, based on empirical results.

The cost function Φ (equation (6.5)) was iteratively minimized using a conjugate gradient (CG) algorithm [HZ05] which was implemented in MATLAB. Therefore, the gradients with respect to S_0 , D , R_2 and Θ were calculated as follows:

$$\begin{aligned} \nabla_{S_0} \Phi(\vec{\rho}, \vec{x}) &= \sum_t \sum_c \operatorname{Re} \left[e^{-b(t)D - t_E(t)R_2 + i\Theta(t)} C_c^* \left(A^\dagger \left(\vec{F}(\vec{x}, t, c) \right) \right) \right] \\ \nabla_D \Phi(\vec{\rho}, \vec{x}) &= \sum_t \sum_c \operatorname{Re} \left[-b(t) S_0^*(\vec{x}, t) C_c^* \left(A^\dagger \left(\vec{F}(\vec{x}, t, c) \right) \right) \right] \\ \nabla_{R_2} \Phi(\vec{\rho}, \vec{x}) &= \sum_t \sum_c \operatorname{Re} \left[-t_E(t) S_0^*(\vec{x}, t) C_c^* \left(A^\dagger \left(\vec{F}(\vec{x}, t, c) \right) \right) \right] \\ \nabla_{\Theta(t_j)} \Phi(\vec{\rho}, \vec{x}) &= \sum_c \operatorname{Re} \left[-i S_0^*(\vec{x}, t_j) C_c^* \left(A^\dagger \left(\vec{F}(\vec{x}, t, c) \right) \right) \right] \end{aligned} \quad (6.7)$$

Here, A^\dagger is the back transformation from Cartesian image space to radial k-space, $(\cdot)^*$ denotes the complex conjugate, and $\operatorname{Re}[\cdot]$ is the real part of the argument. A comprehensive derivation of the formulas (6.7) can be found in [BUF09]. Not shown in equations (6.7) are the additional gradients given by the regularization term in equation (6.5), which are added to each component of $(S_0, D, R_2, \Theta(t_1), \dots, \Theta(t_n))$.

The direction of the gradients is dependent on the input data scaling: In equation (6.7), only the gradient with respect to S_0 does not scale with S_0 , such that an absolute scaling of the signals increases the gradients with respect to D and R_2 , but not S_0 . Additionally, the parameter maps D and R_2 scale with $-b(t)$ and $-t_E(t)$, respectively. Hence, k-space was chosen to be normalized to 1, and the b and t_E values were normalized to a maximum value of 0.08:

$$\begin{aligned} b_{\text{Norm}}(t) &= 0.08 \cdot \frac{b(t)}{\max(b(t))} \\ t_{E,\text{Norm}}(t) &= 0.08 \cdot \frac{t_E(t)}{\max(t_E(t))} \end{aligned} \quad (6.8)$$

These values have been chosen empirically, as they yielded stable convergence for simulated as well as measured data.

During iterations of the CG algorithm, divergence of the phase maps Θ was noticed due to diverging gradients. Therefore, the gradient with respect to Θ (equation (6.7)) was artificially suppressed by a factor of 0.05, while leaving the regularization part of this gradient unchanged, which ensured stable convergence and yielded the physically expected smooth phase maps.

6.2.3 MODEL-INTERLEAVED RECONSTRUCTION

As an alternative to the model-based reconstruction, a model-interleaved scheme was developed [BLB21]. This reconstruction method assumes only the imaging process as a model and reconstructs each image using a conventional compressed sensing algorithm [BUF07]. In contrast to the model-based approach, the function $\vec{F}(\vec{x}, t, c)$ is simplified:

$$\vec{F}(\vec{x}, t, c) = A \left(C_c(\vec{x}) \vec{I}(\vec{x}, t, c) \right) \quad , \quad (6.9)$$

where $\vec{I}(\vec{x})$ is treated as an independent image for each echo time TE. The loss function is given by

$$\Phi = \frac{1}{2} \sum_t \sum_c \|\vec{F}(\vec{x}, t, c) - \vec{y}(t, c)\|_2 + \lambda \|W(\vec{F}(\vec{x}, t, c))\|_1 \quad (6.10)$$

and is iteratively minimized using a conjugate gradient algorithm as before. However, after each iteration in the reconstruction, the model is incorporated by fitting ADC and T₂ to \vec{I} . From the fitted parameter maps, synthetic images \vec{I}' for each echo time are calculated, matched with the phase of the current iteration of \vec{I} . \vec{I}' is then used as the new current estimate for the next iteration, until the algorithm converges.

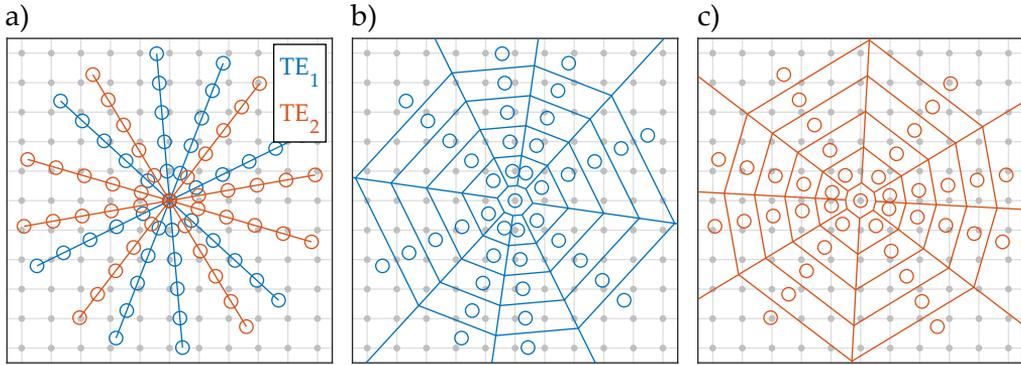


Figure 6.2: K-space trajectories for two TE with 8 spokes each. In a), all spokes are shown. Due to the golden angle sampling, k-space is sampled homogeneously. b) and c) show the trajectories for TE₁ and TE₂, respectively, with their corresponding Voronoi diagrams. The areas in the diagram are the inverse density weighting factors.

6.2.4 DENSITY COMPENSATION

Density compensation is an important processing step of the NUFFT [PM99]: Kaiser-Bessel gridding generally sums signal contributions from different measured points with varying amplitudes which depend on their distance to the sampled point. This sum is independent of the number of measured points which were included, and therefore, a sampling point with a high density of measured points will yield higher values than others. Therefore, signal intensities are weighted with a local sampling density to counter this effect.

There are several ways of calculating a density compensation: in radial imaging with equidistant angles between spokes, the sampling density decreases linearly with the radius. This function is a good estimation for golden angle radial sampling only if the number of acquired spokes is high enough to average out the variable angles. Here, the density compensation is calculated by calculating a Voronoi diagram based on the actual k-space trajectory. The trajectory for two consecutive TEs is shown in figure 6.2 a). In b) and c), the corresponding Voronoi diagrams are shown. The resulting density compensation for each point is given by the inverse of the area and is capped at 1 to prevent overcompensation and edge effects, where the Voronoi diagram has open ends.

6.2.5 ECHO SELECTION FOR RECONSTRUCTION

For the reconstruction, some data should be discarded: To prevent steady state effects in the signal when multiple spokes are acquired, a long repetition time TR is chosen. This choice allows to acquire very long echo trains, with echo

times exceeding 300 ms. However, with increasing diffusion weighting and echo times, the signal contribution exponentially decreases and noise dominates the acquired data. Therefore, echoes with a b-value $> 800 \text{ s/mm}^2$ and $TE > 200 \text{ ms}$ were excluded, as they systematically deviate from the imposed model. This restriction leaves a large dead time between multiple excitations, which can be used to acquire additional slices.

6.2.6 COIL SENSITIVITY MAPS

The coil sensitivity maps were calculated from the raw k-space data using eSPIRIT [Uec14] as implemented in the BART toolbox [Uec21]. To improve accuracy of the sensitivity maps, raw data from different b-TE-pairs are combined. This combination can lead to image artifacts, but substantially improves k-space coverage. Since sensitivity maps are inherently low in resolution, possible frequency artifacts can be tolerated in exchange for better overall quality.

6.2.7 GRADIENT DELAY CORRECTION

Radial acquisition can be sensitive to gradient delay errors, as it leads to a blurring in k-space center, rather than a shift in multiecho Cartesian sampling [Duy98; PDM03; Mou14]. In this work, gradient delays in the k-space trajectory were corrected using the RING method [RHU19] as implemented in the BART toolbox as a pre-processing step.

The RING method estimates time offsets τ for each physical gradient axis based on the measured data. Therefore, it computes the position of the maximum in each spoke of the nominal trajectory. With small gradient delays, these maxima are shifted related to the expected position $k = 0$. RING then fits an ellipse to the calculated maximum positions, which yields a good estimator for the actual delays τ for each axis. Using the gradient delays, a corrected trajectory can be calculated which is used in the reconstruction.

6.3 SIMULATION

Simulations were performed to validate various aspects of the sequence design. Bloch simulations were used to analyze the echo formation and exclude artifacts from stimulated echoes. Simulations based on two different analytical phantoms were conducted to test and optimize the reconstruction: a Shepp-Logan type of phantom [SL74] to test geometric and numeric accuracy, and a simple structured phantom to search for limiting cases of T_2 -ADC combinations.



Figure 6.3: Normalized Shepp-Logan phantom used in the Bloch simulations.

6.3.1 BLOCH SIMULATION

A Shepp-Logan phantom (figure 6.3) was prepared with 360 isochromats in each pixel, a matrix size of 64×64 and a constant $T_1 = 700$ ms and $T_2 = 90$ ms in all pixels. An echo spacing of 12 ms was assumed, and 8 consecutive echoes were simulated, with diffusion blocks before the 1st, 4th and 7th.

Two separate simulations were performed: In the first simulation, dephasing by the diffusion gradients was set to 5 times the dephasing of the crusher gradients. The second simulation used maximal diffusion gradients, which are a factor of 5.2593 greater than the crusher gradient moment and therefore the CPMG condition is violated. Additionally, an imperfect refocusing pulse with a flip angle of 160° was assumed to determine the influence of the imperfect slice profiles.

The pulse sequence was described by a set of vector operations which account for signal relaxation (T_1 and T_2), signal dephasing, including optional off-resonance, and the signal response to RF-pulses. The signal at a given time t can then be calculated by repeated application of the matrices, given an initial signal $M(0)$:

$$\vec{M}(t) = (R \cdot D \cdot P)^n \cdot \vec{M}(0) \quad (6.11)$$

where R , D and P are the matrix formulations for relaxation, dephasing and pulse response derived from the Bloch equations, and n is the echo train length. For one isochromat R , D and P , are given by

$$R = \begin{pmatrix} e^{-t/T_2} & 0 & 0 & 0 \\ 0 & e^{-t/T_2} & 0 & 0 \\ 0 & 0 & e^{-t/T_1} & 1 - e^{-t/T_1} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$D = \begin{pmatrix} \operatorname{Re}(e^{-i\phi}) & 0 & 0 & 0 \\ 0 & \operatorname{Im}(e^{-i\phi}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$P = \begin{pmatrix} \cos(\alpha) \sin^2(\varphi) + \cos^2(\varphi) & \sin^2(\frac{\alpha}{2}) \sin(2\varphi) & -\sin(\alpha) \sin(\varphi) & 0 \\ \sin^2(\frac{\alpha}{2}) \sin(2\varphi) & \cos(\alpha) \cos^2(\varphi) + \sin^2(\varphi) & \cos(\varphi) \sin(\alpha) & 0 \\ \sin(\alpha) \sin(\varphi) & -\cos(\varphi) \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6.12)$$

and the magnetization vector of the isochromat is expressed as

$$\vec{M} = \begin{pmatrix} M_x \\ M_y \\ M_z \\ M^{\text{eq}} \end{pmatrix}, \quad (6.13)$$

with M^{eq} being the equilibrium magnetization given by the Shepp-Logan phantom. The angle ϕ is the dephasing due to the gradient moments, φ is the offset angle of the pulse ($\pi/2$ for the 180° pulses and 0 for the 90° pulse), and α is the flip angle. The MATLAB implementation can be found in chapter 11.2, listing 11.3 in the appendix, and for a similar derivation also see [Rau11]. No dephasing due to molecular diffusion is included in the simulation ($D = 0$). However, since diffusion is a purely statistical signal attenuation which cannot be rephased, it can be omitted in the simulations without an effect on the echo formation. Still, the diffusion gradients were included in the simulation to account for the additional signal dephasing. To separate the the expected artifact from radial sampling artifacts, a Cartesian readout trajectory and a fully sampled k-space were used in the simulation.

6.3.2 ANALYTICAL PHANTOM RECONSTRUCTION

The reconstruction algorithm was tested on numerical phantom data, which was calculated analytically in k-space [KSÖ07]. Raw data was then sampled according to the sequence parameters and the k-space trajectory was taken directly from *Pulseq* after sequence construction. The phantom is defined by a number of ellipses where position and signal amplitudes can be chosen freely. Distinct pairs of ADC and T₂ values were assigned to each ellipse. The absolute

PARAMETER	PHANTOM a)	PHANTOM b)
ΔTE	[11, 11.5, 12] ms	[11, 11.5, 12] ms
#Readouts (Echoes)	15	15
#b-TE-Pairs	45	45
Diffusion block before echo	4, 7, 10, 13	1, 6, 11
Min./Max TE	[11 / 288] ms	[11 / 252] ms
Min./Max b-value	[0 / 831] s/mm ²	[132 / 581] s/mm ²

Table 6.1: Sequence parameters used for the two phantom experiments.

signal amplitudes were calculated according the signal model in equation (6.4) for each echo of the simulated sequence. No additional phase Θ was applied to the phantom. Two different phantom models, shown in figure 6.4, were used to test the reconstruction.

The first model (phantom a)) is similar to the original Shepp-Logan phantom [SL74] to represent realistic anatomical features. There is a large overlap between the two ellipses on the top left with additive mixing signals. To conform to the monoexponential decay for both T_2 and ADC, the overlapping areas were defined using identical T_2 and ADC. Hence, the resulting signal is given by

$$\begin{aligned} S &= S_{0,1} e^{-bD_1 - TE/T_{2,1}} + S_{0,2} e^{-bD_2 - TE/T_{2,2}} \\ &= (S_{0,1} + S_{0,2}) e^{-bD - TE/T_2} \end{aligned} \quad (6.14)$$

$$\text{with } D_1 = D_2 = D, \quad T_{2,1} = T_{2,2} = T_2$$

and thus only the amplitude S_0 is different. There are two more overlapping regions which incorporate mixed T_2 and ADC values and were kept to test reconstruction performance in areas which do not strictly follow a monoexponential model function. This simulates the case of voxels with two major compartments, e. g. on the border between two tissue types.

The second model (phantom b)) consists of a 9×9 grid of circles with combinations of ADC values from 0.2 to $2 \cdot 10^{-3}$ mm²/s and T_2 values from 0 to 200 ms. T_2 increases in left-right direction and ADC increases in top-bottom direction. This model was used to test the reconstruction accuracy for a larger number of different combinations of ADC and T_2 values.

The simulation was run for the set of parameters presented in table 6.1. Results were compared against the analytical ground truth, qualitatively in the form of difference images and quantitatively by a region-based analysis.

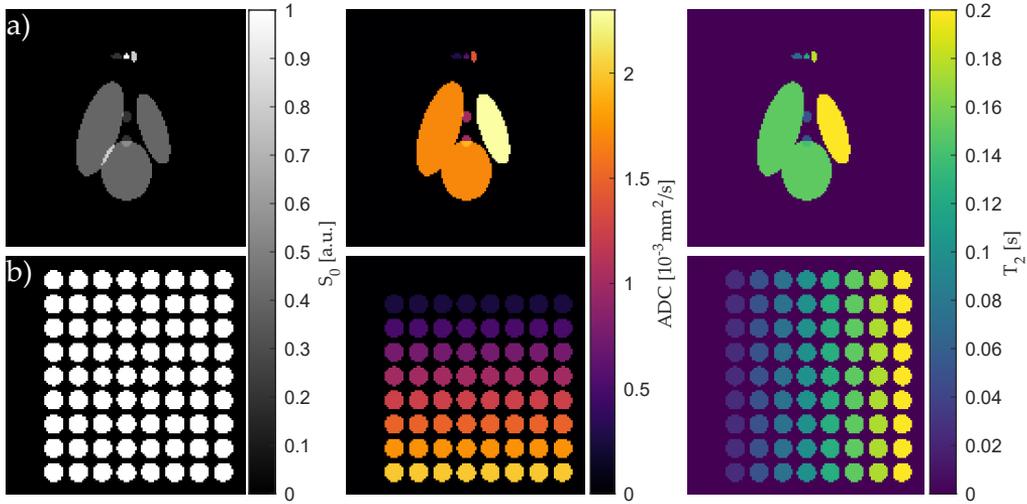


Figure 6.4: Two numerical phantoms used to quantify the SATM reconstruction performance. a) is a Shepp-Logan type of phantom with adjusted signal intensities. The grid in b) is used to sample many different combinations of ADC and T₂.

6.4 MEASUREMENTS

The SATM sequence was implemented as described above for a clinical 3T MRI system (PRISMA, Siemens, Erlangen, Germany) using the sequence prototyping environment *Pulseq* [Lay17; BLB21]. It was tested using two different phantoms, shown in figure 6.5, and, after parameter optimization, also *in vivo*.

The first phantom (measurement phantom a)) is a diffusion phantom consisting of 4 compartments: Three 50 ml tubes filled with a gel that were embedded in a gel-filled bag. The phantom was originally designed for a human breast cancer study and adopted for testing of the sequence.

Measurement phantom b) was designed in association with [Rac21] to test the sequence in different combinations of T₂ and ADC values. It consists of 12 separate tubes of 15 ml each, which are placed in a holder and submerged in a water tank during measurements. The tubes are filled with different concentrations of polyethylenglycol (PEG) and a Gd-based MRI contrast agent (Gadovist, Bayer Vital GmbH, Germany) to independently vary T₂ and ADC in the tubes [Gat14]. ADC and T₂ values of the tubes are listed in table 6.2.

Finally, measurements were performed of the brain of a healthy volunteer. In this early stage of sequence development, the human brain is an optimal target structure for sequence testing. In the head, minimal motion is expected (no breathing motion, tongue motion, heartbeat, etc.), but good ADC and T₂ contrasts are available between gray and white matter and CSF. The robustness

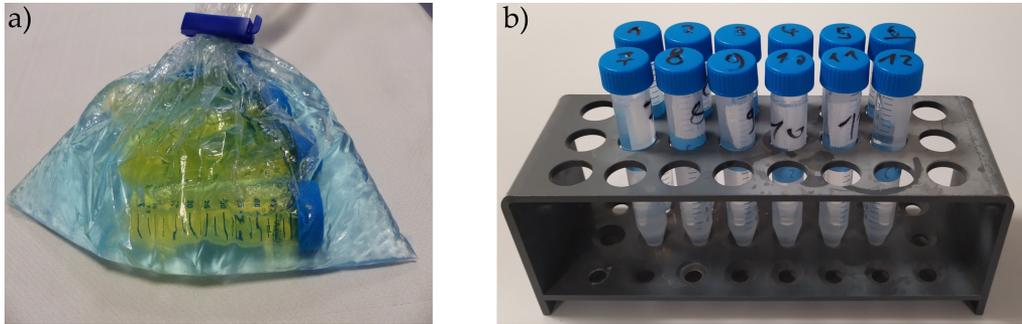


Figure 6.5: Measurement phantoms a), adapted from a human breast diffusion study. Note, that at the time of the measurements only 3 tubes were present in the phantom, a 4th tube was added later. Measurement phantom b) systematically varies T_2 and ADC using PEG and a Gd-based contrast agent (photo reproduced from [Rac21, p. 7, fig. 2]).

TUBE	T_2 [ms]	ADC [mm^2/s]
1	45 ± 10	0.60 ± 0.09
2	55 ± 7	2.06 ± 0.15
3	66 ± 6	1.43 ± 0.15
4	61 ± 7	0.65 ± 0.04
5	89 ± 8	2.10 ± 0.10
6	110 ± 9	0.66 ± 0.04
7	131 ± 9	1.47 ± 0.07
8	127 ± 10	0.86 ± 0.08
9	156 ± 14	0.64 ± 0.10
10	209 ± 18	2.04 ± 0.03
11	226 ± 19	0.87 ± 0.07
12	264 ± 20	1.46 ± 0.06

Table 6.2: Reference values for the measurements phantom b), taken from [Rac21].

to distortion artifact can be tested in areas close to the paranasal sinuses which can cause strong susceptibility artifacts in conventional DW-EPI measurements.

For both phantoms and *in vivo*, T_2 and ADC reference measurements were performed in combination with the proposed sequence, which allowed a pixel-wise and a region based comparison of the reconstructed maps. The reference T_2 map was acquired using a FSE with up to 32 echoes and TEs between 13-442 ms. The ADC map was acquired using a conventional DW-EPI with b-values of [50, 400, 800] s/mm². Matrix size, FOV and slice thickness were kept identical between reference and new sequence. For the volunteer experiment, an additional high-resolution T_2 -weighted FSE image was acquired for anatomical reference, with the same FOV but different matrix size.

Part III

RESULTS

AUTOMATED TUMOR SEGMENTATION

7.1 SINGLE SHOT EPI VS. READOUT SEGMENTED EPI

The FMISO trial MRI protocol included both, ssEPI and rsEPI diffusion sequences. Data from a ssEPI sequence was analyzed with respect to the apparent Nyquist ghost artifact. Figure 7.1 shows two slices from the same acquisition, one with a 4-fold ghost artifact, and the other without. The plots show the corresponding three navigator lines which are acquired before each EPI readout. For the case without any apparent ghost artifact, the navigators show a single peak, while the corrupted slice shows a series of multiple peaks that are hard to differentiate. The failing echo correction can then lead to the ghosting artifact as described in section 3.2.2.

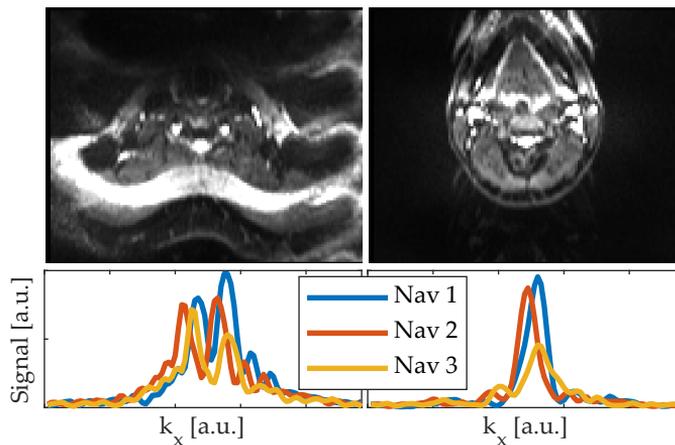


Figure 7.1: Upper row: Two slices of a DW-ssEPI sequence ($b = 50 \text{ s/mm}^2$). The left slice shows the lower part of the neck, and a 4-fold Nyquist ghost artifact is visible. The fat signal is clearly visible despite a spectral selective fat saturation pulse. The right image from the same measurement shows a more cranial slice. No Nyquist ghost artifact is visible. The bottom row shows the corresponding navigator lines, which are acquired with each image for echo correction.

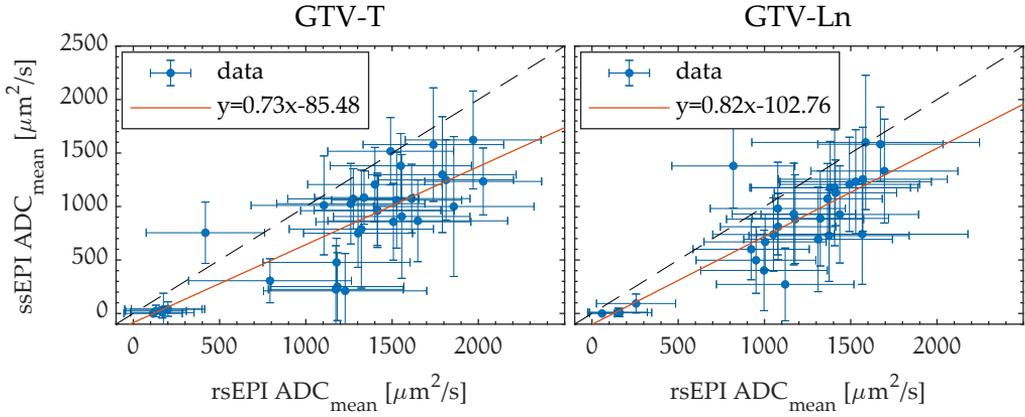


Figure 7.2: Comparison of ADC values for ssEPI and rsEPI in tumor (GTV-T) and lymph node metastases (GTV-Ln). The dashed gray line is the line of identity.

The clinical trial shows that rsEPI is generally less corrupted by the Nyquist ghost artifact than ssEPI: Out of 76 measurements with rsEPI, 61 (80%) were found without artifact, and 8 (10%) showed severe artifact throughout the image, such that they could not be used clinically. The remaining 7 images showed some artifacts, but could still be used because the artifacts did not obscure the clinically relevant areas. For ssEPI, only 7 (10%) out of 67 measurements showed good quality without the artifact, while 44 (66%) were clinically unacceptable.

The two sequences were then compared in terms of ADC values in tumor and lymph node tissue, as defined by the ground truth regions GTV-T and GTV-Ln drawn by clinical experts on the superposition of all available MR images. In the tumor, rsEPI yielded significantly higher ADC values than ssEPI, with an average difference of $420 \pm 316 \mu\text{m}^2/\text{s}$. Similarly, for lymph node metastases, the difference was $303 \pm 265 \mu\text{m}^2/\text{s}$. Figure 7.2 shows the differences for each patient in a scatter plot, revealing the systematic nature of the differences between the two sequences. A linear regression yields

$$\begin{aligned} \text{ADC}_{\text{ss},\text{T}} &= 0.73 \text{ADC}_{\text{rs},\text{T}} - 85.48 \mu\text{m}^2/\text{s} \\ \text{ADC}_{\text{ss},\text{Ln}} &= 0.82 \text{ADC}_{\text{rs},\text{Ln}} - 102.76 \mu\text{m}^2/\text{s} \end{aligned} \quad (7.1)$$

for tumor and lymph node metastases, respectively.

7.2 B₀ FIELD HOMOGENIZATION

Field map measurements on two volunteers were acquired to quantitatively and qualitatively assess B₀ field distributions in the head and neck area. The effect of an additional pillow for field homogenization was analyzed. Figure 7.3

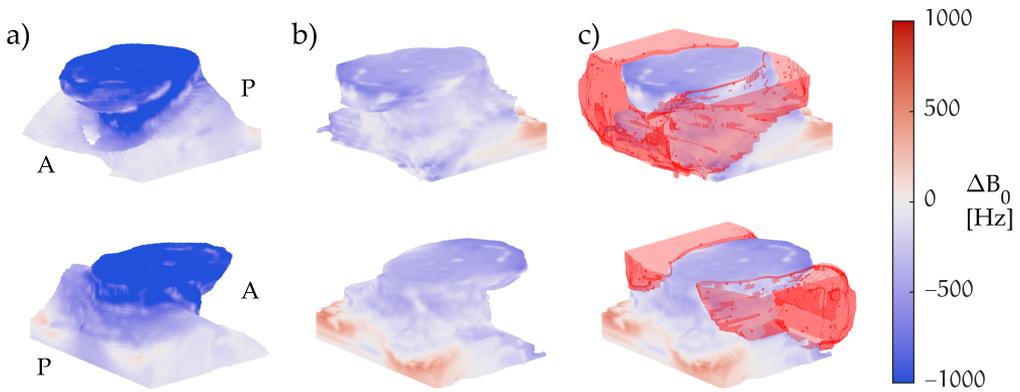


Figure 7.3: 3D rendering of the B₀ field distribution in the head and neck area of a volunteer. Measurements with and without a susceptibility matching pillow are shown from two different perspectives (rows). a) without pillow. b) with pillow. c) same as b) but with the pillow included in the rendering. For the data of a second volunteer, refer to figure 11.1 in the appendix.

shows a ΔB_0 -map for a volunteer with and without an additional susceptibility-matching pillow. A large B₀ difference of 850 Hz is visible in the head-foot direction (figure 7.3 a), which was reduced to 350 Hz when the pillow was included (figures 7.3 b), c) and 7.4 a)).

Additionally, the B₀ homogeneity within each slice was significantly improved. The mean $\mu(\Delta B_0)$ and standard deviation $\sigma(\Delta B_0)$ were used to quantify the homogeneity within the imaging volume, and were calculated only within the body. Without the pillow, the mean of $\sigma(\Delta B_0)$ across all 30 slices was $\overline{\sigma(\Delta B_0)} = 220$ Hz. When the pillow was added, this value decreased to $\overline{\sigma(\Delta B_0)} = 141$ Hz, improving homogeneity by 36%.

Next, the setup was adapted to account for the thermoplastic fixation mask and the patients sensibility to any additional load on the neck area, by replacing the U-shaped pillow with two separate water bags placed on each side of the neck. Fieldmap measurements in a volunteer showed a similar effect using two water bags placed on each side of the head. Figure 7.5 shows representative slices from the measurement. The B₀ variation in head-foot direction is now decreased by 38% from 362 Hz to 261 Hz. Similarly, the B₀ variation within each slice was reduced by 17% (see figure 7.4 b)).

For the patient study the amount of the Nyquist ghost artifact present in the images was evaluated. A total of 47/54 image volumes without water bags and 20/22 with water bags were included for ssEPI/rsEPI respectively. The results are presented in table 7.1 and show a clear reduction of artifact given the new setup. With an increase of 21%, an overall very good image quality for 95% of the cases could be reached with rsEPI. For ssEPI, the number of clinically

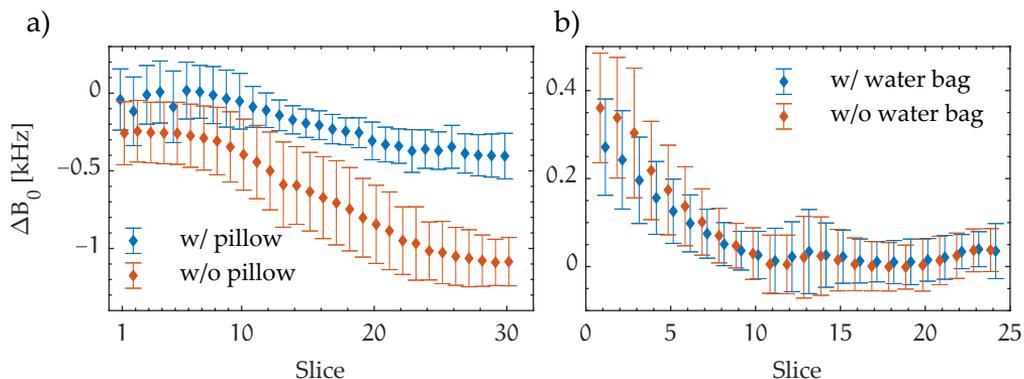


Figure 7.4: Mean $\mu(\Delta B_0)$ (points) and standard deviation $\sigma(\Delta B_0)$ (error bars) for two experiments in two volunteers. $\mu(\Delta B_0)$ and $\sigma(\Delta B_0)$ were measured with and without a U-shaped pillow on top of the neck in a). b) shows data from a different volunteer, who was imaged with two water bags to each side of the neck, in the same setup as patients in the FMISO trial.

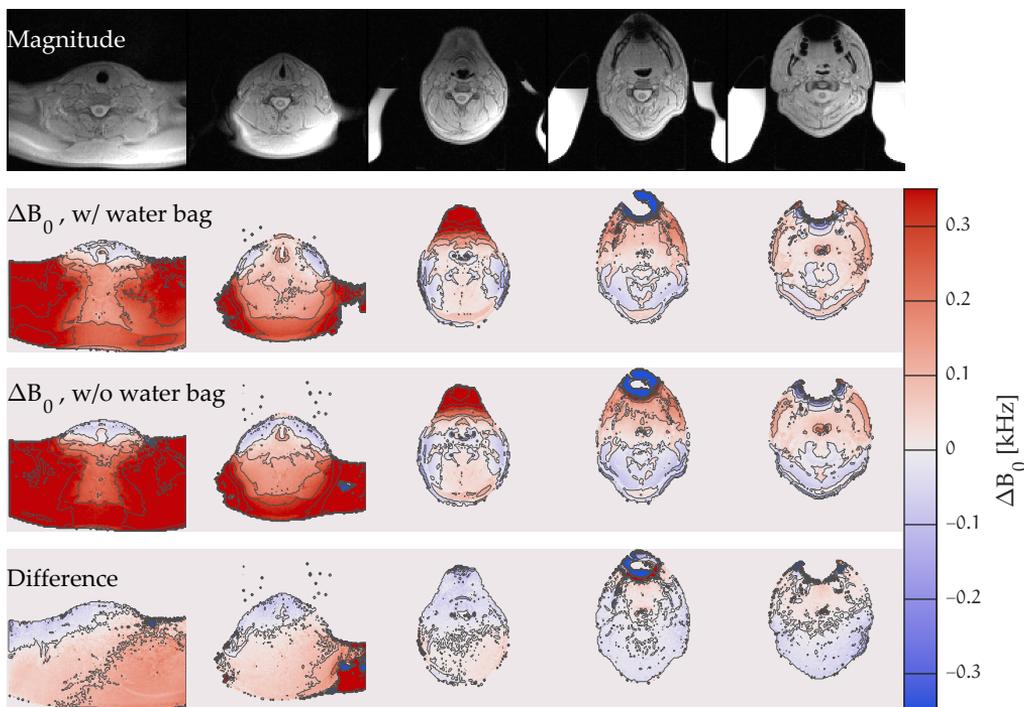


Figure 7.5: Field map measurements in the volunteer experiment. The first row shows magnitude images of the gradient echo fieldmap sequence, for reference. The second and third rows show the ΔB_0 -maps which were acquired with and without the water bags to each side of the head. The last row shows the difference image of the two cases.

GRADE	SSEPI		RSEPI	
	W/O WATER BAG	W/ WATER BAG	W/O WATER BAG	W/ WATER BAG
0	40 (85%)	4 (20%)	8 (15%)	0 (0%)
1	5 (11%)	11 (55%)	6 (11%)	1 (5%)
2	2 (4%)	5 (25%)	40 (74%)	21 (95%)
Total	47	20	54	22

Table 7.1: Analysis of the Nyquist ghost artifact frequency in the FMISO trial in the original setup (w/o water bag) and the new setup (w/ water bag).

unusable images could be reduced from 85% to 20%, while still only 25% of the images were free of artifact. In terms of usability, the new setup was not perceived as uncomfortable by the patients, and the arrangement of the water bags was fast, not significantly increasing the overall examination time in the MRI for the patients.

7.3 DISTORTION CORRECTION

7.3.1 RETROSPECTIVE DISTORTION CORRECTION

To test the distortion correction algorithm, fieldmaps data were acquired in a volunteer experiment in the head and neck area (figure 7.6) [Bie19b]. Anatomically precise T_2 weighted FSE data is shown in green, while the purple overlay shows DW-EPI data ($b = 50 \text{ s/mm}^2$). Further, the respective correction fields are shown in figure 7.6 a), which are calculated either from the measured fieldmap (left) or from the proposed correction algorithm by comparison of T_2 -w and EPI images. Generally, the measured fieldmap is considered ground truth. The computed map from the correction algorithm reproduces the ground truth map well, with an Euclidean image distance D^2

$$D^2 = \frac{1}{N} \sum_{x=1}^N (I_1(x)^2 - I_2(x)^2)$$

between the two maps $I_1(x)$ and $I_2(x)$ of 5.4 px^2 . Here, x are all pixels inside the neck area. For some patients, the B_0 offresonance leads to a field gradient above 1 pixel shift per image pixel, that retrospective correction becomes impossible. This happens especially at tissue-air boundaries, as seen in figure 7.6 a) around the trachea. Here, both distortion correction algorithms fail. Since the artifact affected distorted and undistorted images similarly, no difference due to the

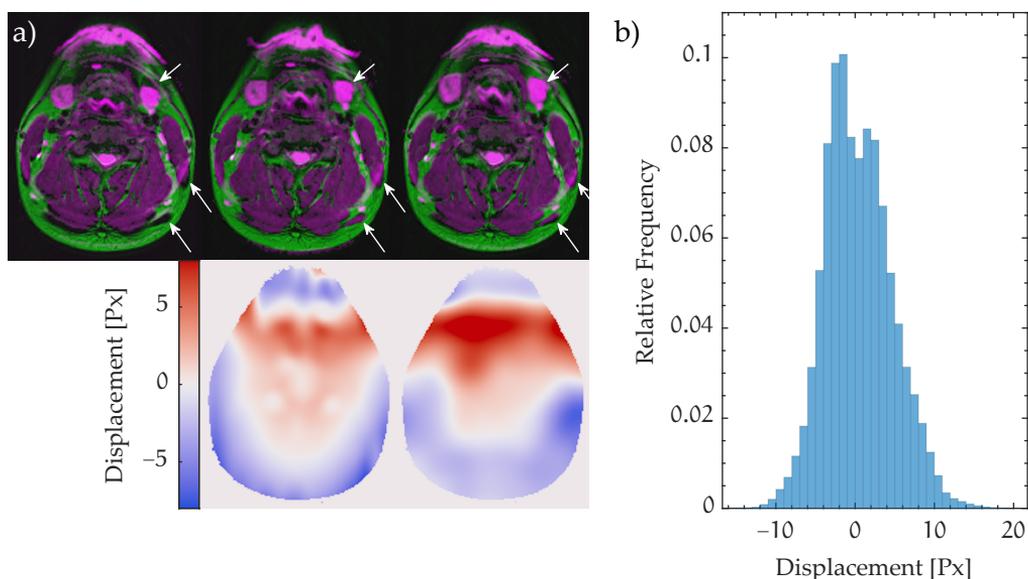


Figure 7.6: a) T₂-weighted (green) and DW-EPI (magenta) images as an overlay. Left: the original DW-EPI, middle: the distortion corrected DW-EPI using the measured field map (bottom row, left). Right: the distortion corrected DW-EPI using the new algorithm, with the corresponding correction maps in the bottom row (right). b) shows the histogram of all pixels of correction maps across all patients, with a mean and standard deviation of 0.46 ± 4.24 pixels. Figure adapted from [Bie19b, fig. 1].

correction are expected in the comparison of segmentation performance. Severe misalignments in the uncorrected data are seen (white arrows), which are corrected using either correction method.

The validated algorithm was then applied to all available DWI data. Figure 7.6 b) shows the relative amount of distortion across all patients in a histogram, as calculated from the distortion correction maps. The mean and standard deviation are 0.46 ± 4.24 pixels. The standard deviation of > 4 pixels confirms the presence of strong distortion across the whole dataset.

7.3.2 SEGMENTATION RESULTS

In this experiment, 28 complete CNNs were trained from scratch as described in section 5.5.5. The training showed a progression up until the 20th epoch, when the validation curve flattens out (see figure 5.12). During training, the validation data of the distortion corrected CNN yields higher sensitivity than the uncorrected CNN. Figure 7.6 shows the segmentation result (red) for a patient trained on data with and without distortion correction. The DSCs were 0.59 and 0.40 respectively. Both segmentations correctly located the tumor,

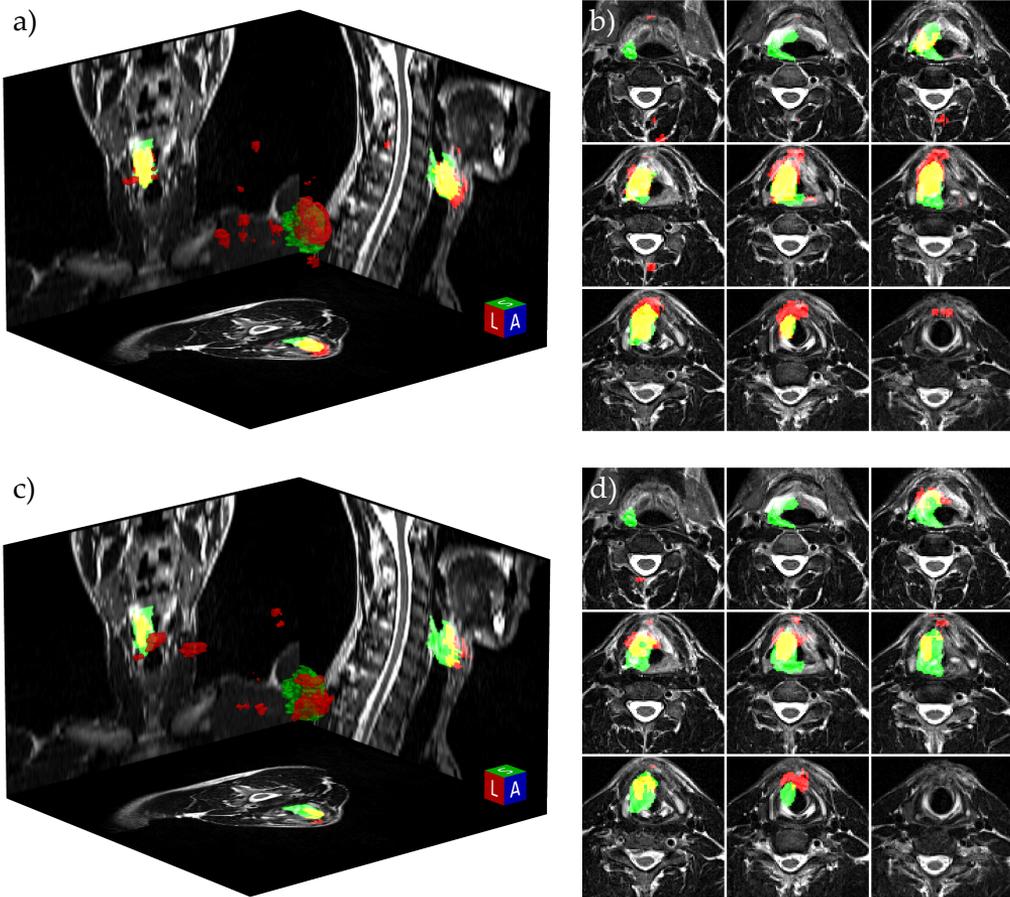


Figure 7.7: 3D visualization of a tumor segmentation. The green regions are the ground truth segmentations and red regions are predicted by the CNN. a) and b) show the results from a CNN trained on distortion corrected DWI, and c) and d) are based on uncorrected DWI. Reproduced from [Bier19b, fig. 3].

but the distortion corrected segmentation shows significantly higher overlap between the predicted and the ground truth segmentation. Additionally, both CNNs segmented smaller regions far from the ground truth as seen in the 3D visualization, which could easily be identified as false positives by a trained viewer.

Across all test patients, the segmentation performance reached DSCs of up to 0.68/0.65 for the corrected/uncorrected sets respectively. On average, the segmentation performance were 0.40 ± 0.18 and 0.37 ± 0.21 respectively, but the improvement of using distortion corrected ADC for training could not be shown to a statistical significant degree ($p = 0.313$). Figure 7.8 shows the reported results in terms of the DSC for all test patients in a scatter plot.

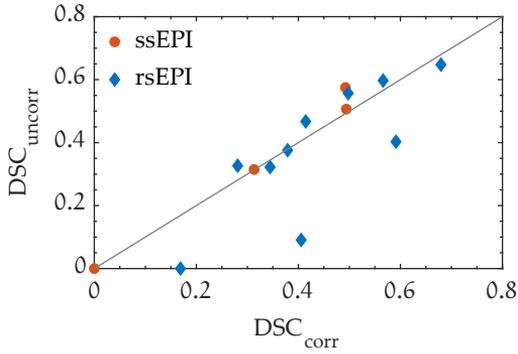


Figure 7.8: Results from the 14-fold leave-one-out test strategy. Most segmentations are close to the line of identity (solid gray line), while for 3 cases the distortion corrected segmentation performed significantly better. Reproduced from [Bie19b, fig. 4].

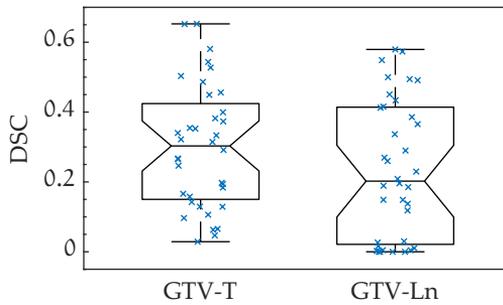


Figure 7.9: Box plot of the reference CNN for all 36 test patients for tumor (GTV-T) and lymph node metastases (GTV-Ln). Figure reproduced from [Bie20b, fig. 2].

7.4 CHANNEL-WISE IMPORTANCE ANALYSIS

This experiment analyzes the relative information content of each of the 7 input channels (T_1 -w, T_2 -w, T_1 -w contrast enhanced (CE), K^{trans} , v_e , ADC and T_2^*) on the performance of a CNN trained to segment GTV-T and GTV-Ln [Bie19a; Bie20b]. The general performance of a network trained on all available channels (reference CNN) is shown in figure 7.9. The best segmentations score 65% / 58% DSC in GTV-T and GTV-Ln, respectively.

Half of the test patients score lower than 30%/20% DSC respectively for GTV-T/GTV-Ln. Figure 7.10 indicates a correlation between tumor size and segmentation performance, with lower performances for smaller lesions. Hence, the large number of low-scoring test patients is possibly linked to very small lesions. This effect is more pronounced in the tumor, possibly because primary tumors are typically singular lesions, whereas lymph node metastases can include multiple smaller regions which together form a larger volume.

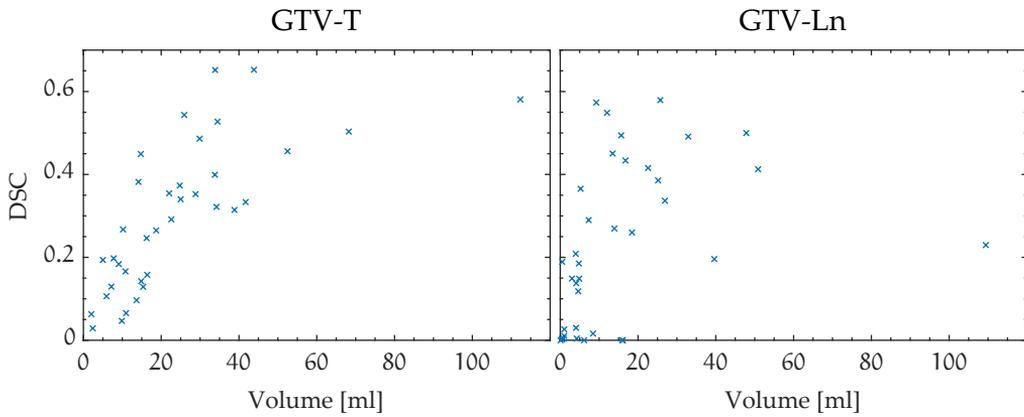


Figure 7.10: Comparison of tumor size and segmentation performance. Very low segmentation performances are restricted to small tumor sizes. Figure reproduced from [Bie20b, fig. 3].

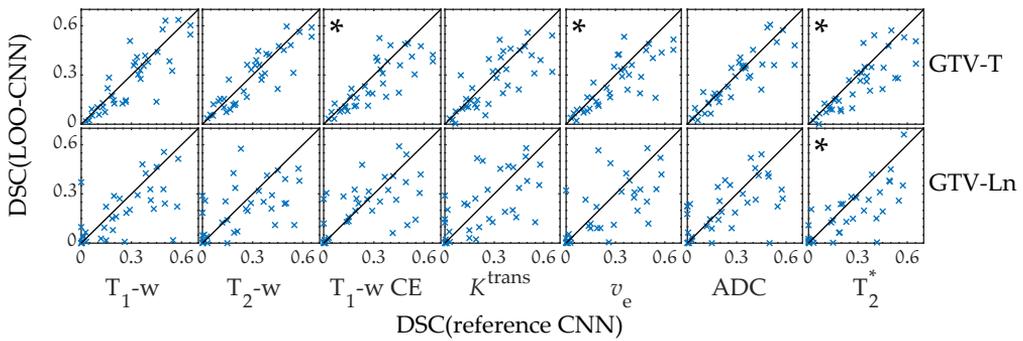


Figure 7.11: Segmentation results from all 288 CNNs. The x-axis always shows the results from the reference CNN, while the y-axis shows the LOO-CNN indicated below. Paired Students t-tests were performed on the results. Significantly differing ($p < 0.05$) distributions are marked by an asterisk. Figure reproduced from [Bie20b, fig. 4].

Next, the relative performance of the networks with single channels left out is compared. Figure 7.11 shows the segmentation results of all 288 trained networks as a scatter plot against the performance of the reference network. The performance of the networks varies between 0% and 65% DSC for all cases. On average, the reference CNN scored higher in each case, as shown in figure 7.12. However, a paired Students-t-test yielded significant deviations ($p < 0.05$) only for segmentation performance between reference CNN and T_1 -w CE in GTV-T, v_e in GTV-T and T_2^* in GTV-T and GTV-Ln (these cases are marked with an asterisk). The overall largest influence on the segmentation performance was found for the T_2^* contrast, with a mean difference in DSC of 0.06 in both GTV-T and GTV-Ln segmentation. An example of the segmentations

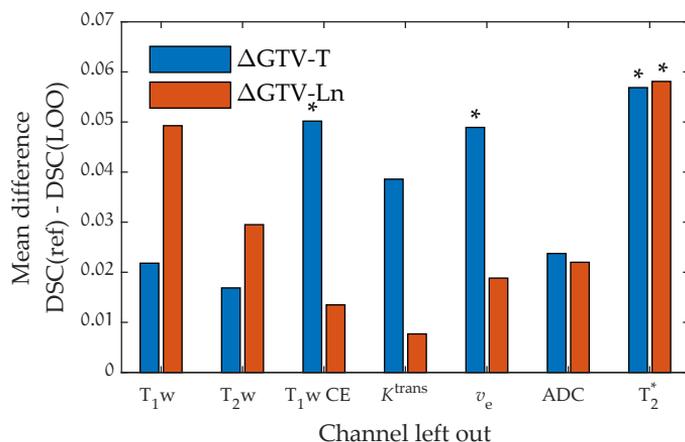


Figure 7.12: Mean segmentation differences between LOO-CNNs and reference CNN. All LOO-CNNs show reduced performance as compared to the reference, but only $T_1w\ CE$, v_e and T_2^* channels yield statistically significant results. Figure reproduced from [Biezob, fig. 5].

of the reference CNN and the T_2^* -LOO-CNN is shown in figure 7.13. The CNN segmentations are marked in blue and orange for GTV-T and GTV-Ln respectively, while the ground truth is shown as a green line. Although in both cases generally too much tissue is segmented, the T_2^* -LOO-CNN shows far greater oversegmentation even in implausible areas like the tip of the tongue. This type of error is observed throughout the test patients, which leads to the decrease in DSC, whereas undersegmentation is much less observed.

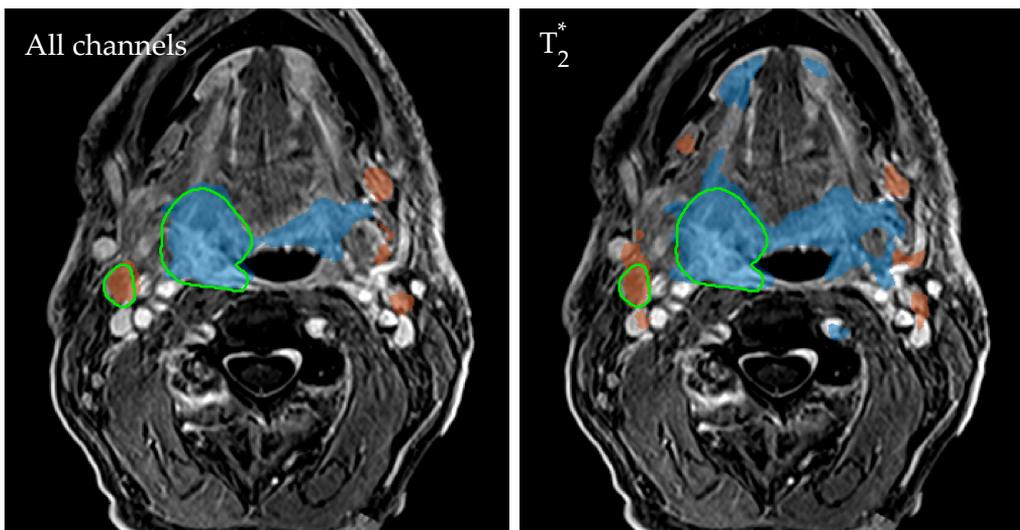


Figure 7.13: Segmentation examples for the reference CNN (left) and the T_2^* -LOO-CNN (right). Figure reproduced from [Bie20b, fig. 6].

SIMULTANEOUS ADC AND T_2 MAPPING

8.1 SIMULATION

The SATM sequence and the reconstruction algorithm described in chapter 6 were validated in Bloch simulations and were tested on analytical phantom data. Here, the simulation results are presented and the reconstruction performance is evaluated for two different phantoms.

8.1.1 BLOCH SIMULATION

Initially, the SATM sequence utilized the maximum gradient strength ($G_{\max} = 80 \text{ mT/m}$) for diffusion sensitization to realize high b-values in minimal time. By violating the second CPMG condition [BKZ04, p. 786], this can lead to stimulated echoes after the second diffusion block which can refocus at times different from the primary spin echoes during readout. Consequently, severe image artifacts can arise, such as ghosting or poor SNR. Bipolar diffusion gradients could circumvent this problem, as their net phase accumulation within one TR vanishes. Unfortunately, at $\Delta TE < 15 \text{ ms}$, bipolar diffusion gradients cannot reach sufficient b-values for diffusion imaging with G_{\max} . Introducing the soft CPMG condition (equation (6.2)), all stimulated echoes are refocused at the same time as the main echo pathway. A phase graph simulation shows the echo formation for violated and soft CPMG condition (figure 8.1).

Bloch simulations were conducted on the sequence, with and without the soft CPMG condition. The simulation included 8 echoes with diffusion blocks after excitation and after echoes 3 and 7. T_1 and T_2 assumed to be constant 700/90 ms for the whole phantom, and the ΔTE set to 12 ms. Figure 8.2 shows the results of the simulations.

The Bloch simulation shows high frequency signal variation artifacts in phase encoding direction (here: left-right direction). If the soft CPMG condition is

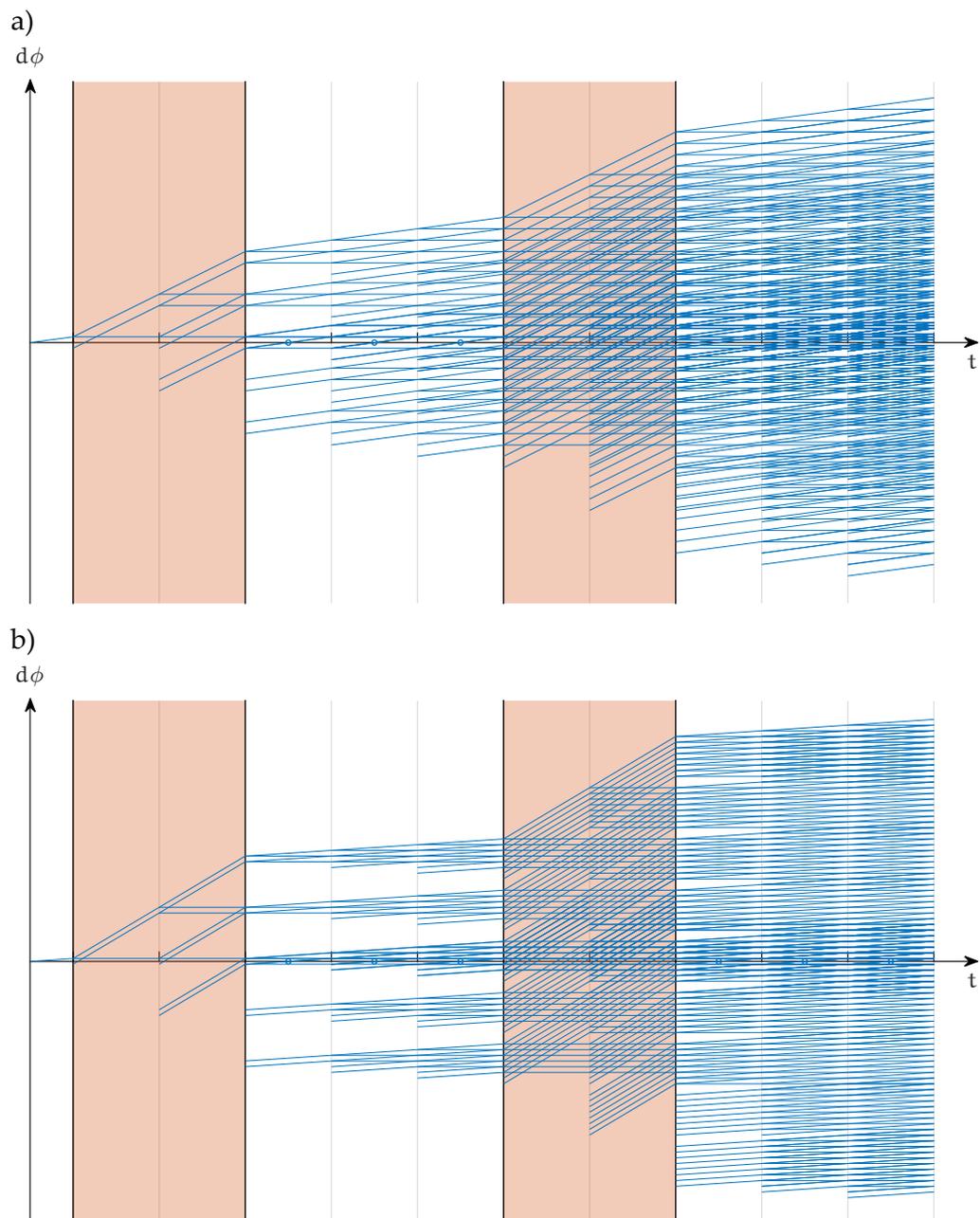


Figure 8.1: Phase graph for the first 10 ΔTE of the SATM sequence. The red boxes mark the diffusion sensitization. a) shows an example with arbitrarily set diffusion gradient strengths: After the second diffusion block, different echo paths refocus at different times during the readout blocks, leading to image corruption. b) shows an example with a diffusion gradient which is a multiple integer of the dephasing due to the crusher gradients: Here, all echoes always refocus exactly in the center of each echo block.

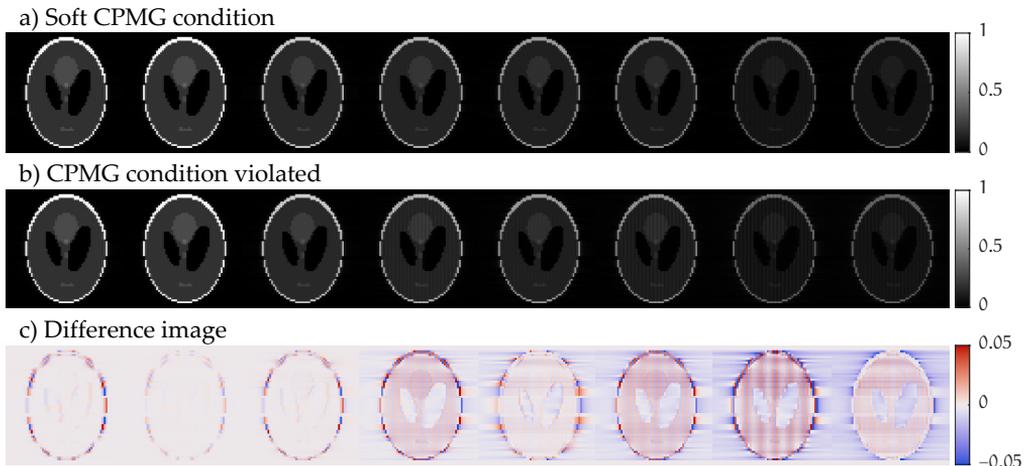


Figure 8.2: Results of the Bloch simulation. In a) the moment of the diffusion blocks was five times the moment of the crusher gradients. In b), the moment was slightly higher, with a factor of 5.2593 the crusher moment, simulating the maximal gradient amplitude of the PRISMA MRI system. c) shows the difference between a) and b).

met, the artifact is faintly visible in echoes 7 and 8, the remaining echoes are artifact free. When the CPMG condition is violated, pronounced artifacts can be seen from echo 4 onward. The difference images show a systematic signal variation of up to 5%.

8.1.2 ANALYTICAL PHANTOM RECONSTRUCTION

Numerical phantom data were simulated as described in section 6.3.2. In a first numerical experiment [BLB21], a phantom a) was reconstructed from k-space data with varying numbers of spokes between 1 and 50. The reconstruction was performed according to the interleaved model within 10 iterations. Figure 8.3 shows the ground truth (top left) and the corresponding reconstructions. With 11 or more spokes, no apparent artifact can be seen in the reconstruction anymore. Reconstruction with less than 11 spokes could lead to excessive noise, streaking artifacts, strong edge enhancement effects leading to a complete signal loss at 2 or 1 spokes. The experiment was repeated with complex Gaussian noise added to the k-space data. The noise amplitude was chosen such that a fully sampled, Cartesian trajectory image with otherwise identical parameters could be reconstructed to an image with an SNR of 50 in the top left large ellipse. This noise amplitude was then used for all reconstructions on different numbers of spokes. In this way, the results are comparable in terms of noise robustness, since in an imaging experiment the noise on each spoke is also

expected to be independent on the total number of acquired spokes. The right side of figure 8.3 shows the reconstruction results for the SNR = 50 case, with a similar general behavior as the noiseless case.

A region-based analysis of the mean and standard deviation of the signal was performed in the top left large ellipse in figure 8.3. The analysis included an additional reconstruction with a SNR of 25 (figure 8.4). In the noiseless case (SNR = ∞ , images with less than 0.85% deviation from the ADC and T_2 ground truth can be achieved with 11 spokes or more. The ADC error increases to 2.8%/12.6% for SNR 50/25, and 1.9%/2.2% for T_2 . From the standard deviation in the region of interest, the lower limit for ADC is given at [0.10, 0.27, 0.40] $10^{-3} \text{mm}^2/\text{s}$ for SNR of [∞ , 50, 25]. For T_2 , the lower limit is [5.2, 18.9, 56.0] ms respectively.

To compare the performance of the two different reconstruction algorithms, the numerical phantom b) presented in 6.4 was used. K-space data were simulated for a sequence with 30 spokes per unique b-TE-pair, resulting in 1350 spokes. The reconstruction results as well as the ground truth are shown in figure 8.5. Generally, the T_2 estimation yielded more accurate results than the ADC estimation across all possible combinations of T_2 and ADC, with mean relative errors of 12% for T_2 and 50% for ADC in the interleaved reconstruction model. Moreover, the interleaved reconstruction model resulted in more accurate results for most combinations of ADC/ T_2 for both, T_2 and ADC, compared to the iterative model (figure 8.5), where the mean relative errors were 229%/60% for T_2 /ADC respectively. In the S_0 images, the iterative model reconstruction shows a distinct checkerboard like artifact. This artifact is not present in the interleaved model, although some Gibbs ringing can be observed. Both reconstruction models fail to correctly reconstruct parts with $T_2=25$ ms (first column). The precision of the reconstructed ADC values varies with the underlying T_2 , reaching an average error of less than $1.20 \cdot 10^{-4}$ / $1.72 \cdot 10^{-4}$ mm^2/s for T_2 values >100 ms for the interleaved and iterative reconstruction models respectively.

8.2 MEASUREMENTS

First, phantom measurements were performed in a diffusion phantom with 4 different compartments [BLB21]. The reconstruction was performed using the interleaved model with TV regularization. Reference maps were acquired and compared to the parameter maps obtained with the SATM sequence. Figure 8.6 shows the resulting images and the difference maps. The last row of figure 8.6 shows the 9th echo of the reference FSE sequence, at TE = 142 ms. The phantom image was acquired with the sequence using 30 spokes per TE,

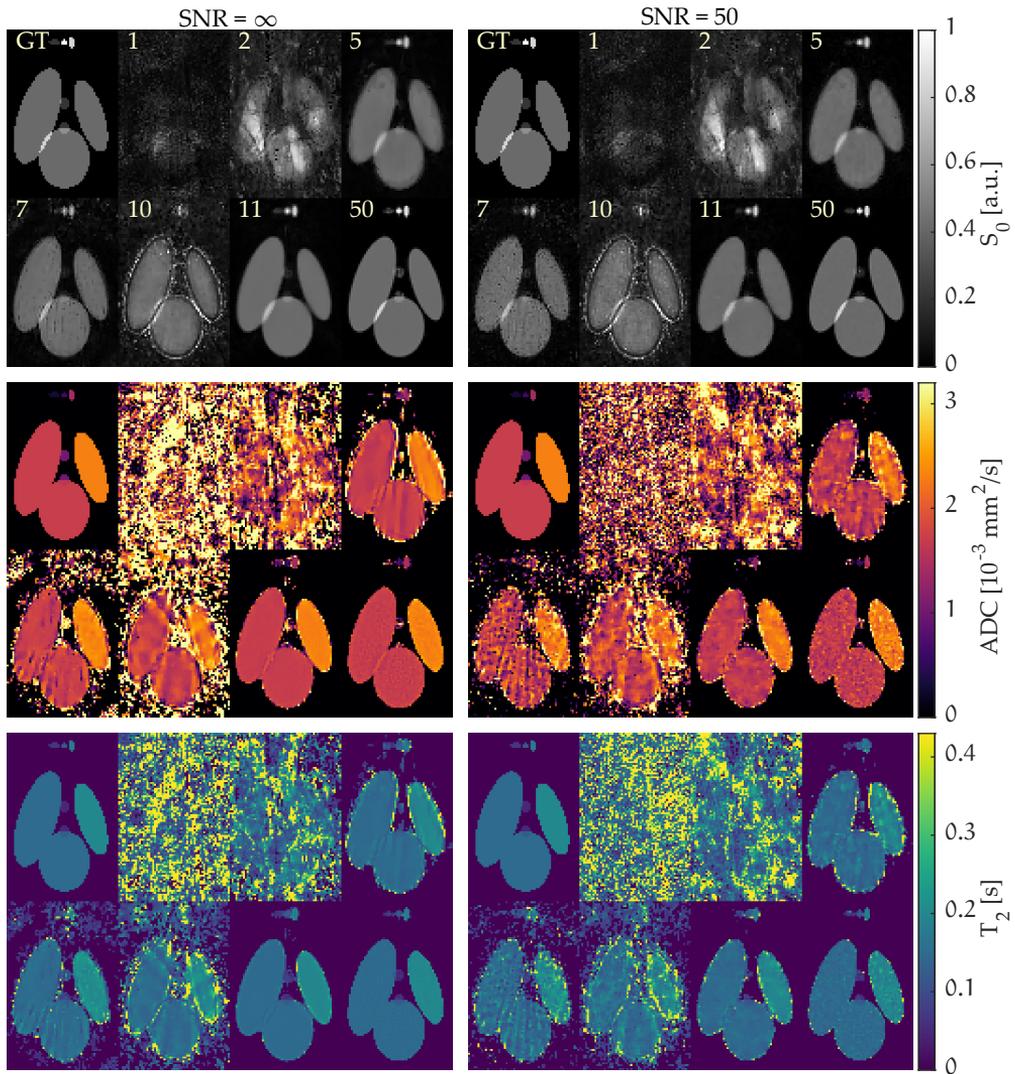


Figure 8.3: Numerical phantom simulations to test the interleaved reconstruction with varying amounts of simulated spokes at two different noise levels. Figure reproduced from [BLB21, fig. 2–3].

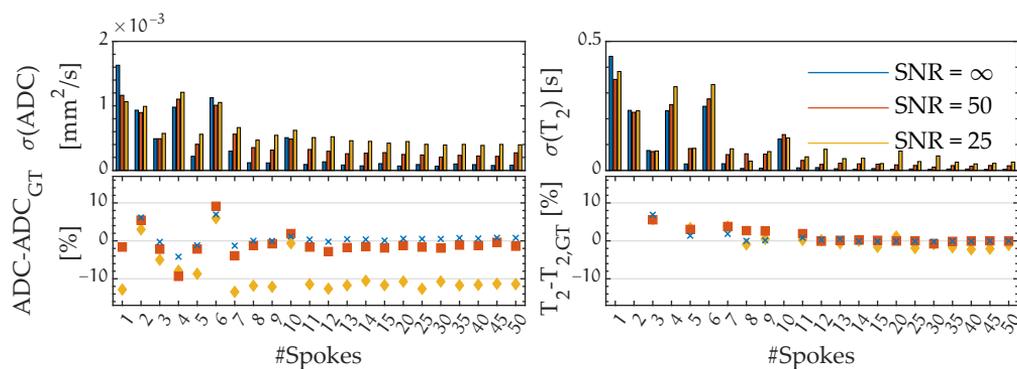


Figure 8.4: Region-based analysis of the reconstruction performance with varying spokes and noise levels in the top left ellipse in figure 8.3. Figure reproduced from [BLB21, fig. 4].

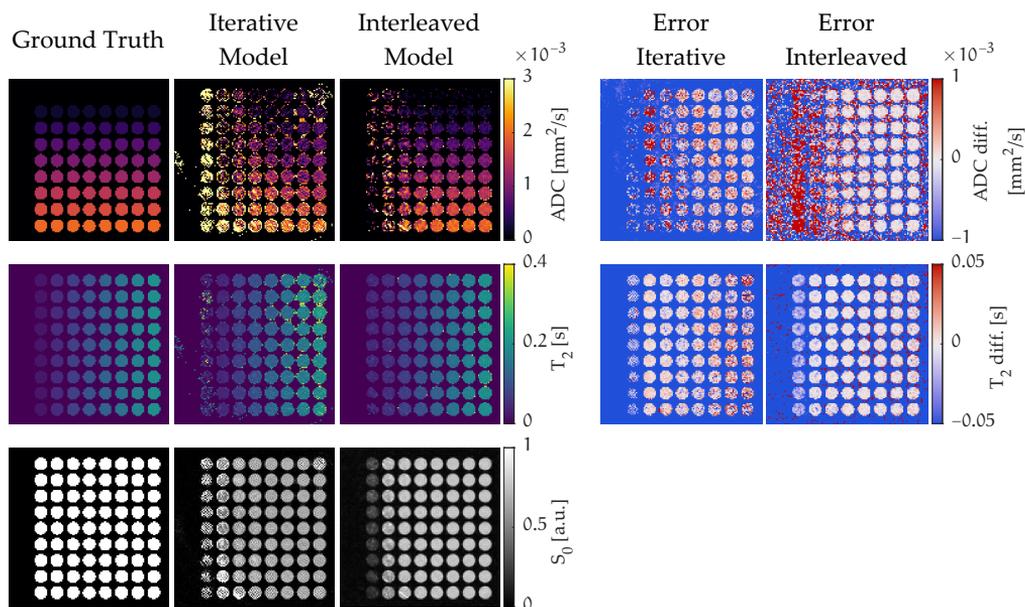


Figure 8.5: Numerical phantom simulations to test the reconstruction performance for different combinations of T₂ and ADC, as well as for the two reconstruction methods, interleaved or iterative. In the phantom, the T₂ value increases from left to right, the ADC value increases from top to bottom.

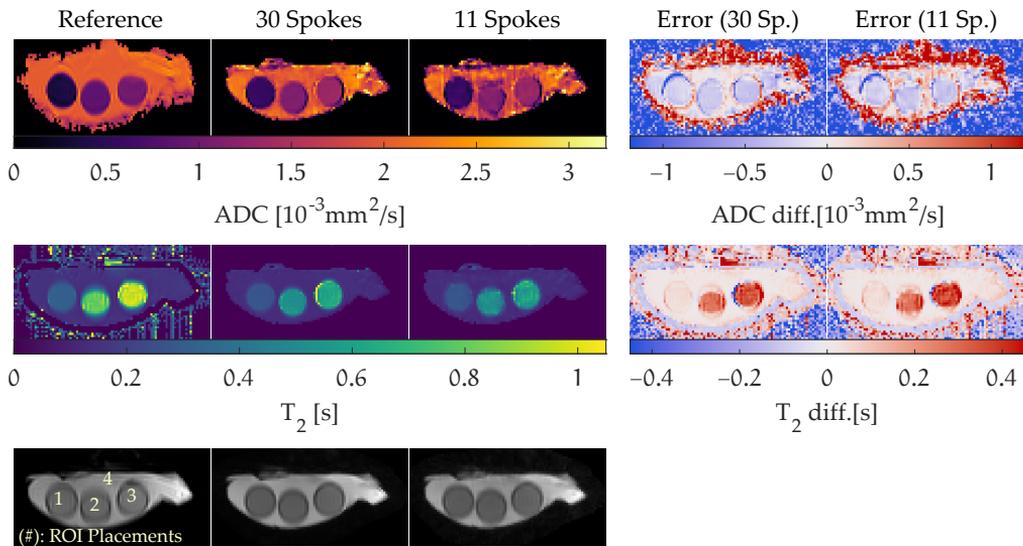


Figure 8.6: Phantom experiment with the new sequence. The phantom consists of 3 vials which are placed in a gel pack. Reference images are taken with a multiecho spin echo sequence and a DW-ssEPI sequence. Figure adapted from [BLB21, fig. 5].

and a ΔTE of [11, 11.5, 12] ms. Two reconstructions are shown, one with the full data and another reconstructed on only 11 spokes per TE. Comparison of the reconstruction with the reference FSE, no geometric distortion can be seen. However, no T_2 related blurring can be seen, as in the reference sequence in phase encoding direction (top-bottom direction). Compared to the reference ADC, an immediate improvement in geometric accuracy is observable. The reconstruction on 11 spokes shows distinctively more noise in the ADC and T_2 maps, but the reconstruction still yields acceptable results.

Figure 8.7 shows the results of the region-based analysis, with the location of the ROIs marked in figure 8.6. Generally, the ADC values show a systematic difference of up to $-0.26 \cdot 10^{-3} \text{ mm}^2/\text{s}$ for ADC values $< 1 \cdot 10^{-3} \text{ mm}^2/\text{s}$ (i. e., 25%). The signal variation is largest for ROI 4, which is outside the tubes, with an ADC of $2 \cdot 10^{-3} \text{ mm}^2/\text{s}$, while it shows the best agreement with the reference measurement. Similarly, ROI 4 agrees best in terms of T_2 with the reference, while ROIs 2 and 3 show a deviation of up to 0.34 s.

In association with [Rac21], a second phantom was designed with different ADC and T_2 combinations, comparable to those found in patients: T_2 ranged from 45-264 ms and ADC ranged from 0.6-2.1 mm^2/s . Using the phantom, ΔTE , the diffusion dead time and the frequency of diffusion sensitization blocks were optimized. As a result, an optimized protocol for an in vivo experiment was defined with the following parameters: $TR = 5 \text{ s}$, $\Delta TE = 9, 9.5, 10, 12 \text{ ms}$, no diffusion dead time (0 s), and diffusion blocks before the 1st, 6th, 11th and 16th

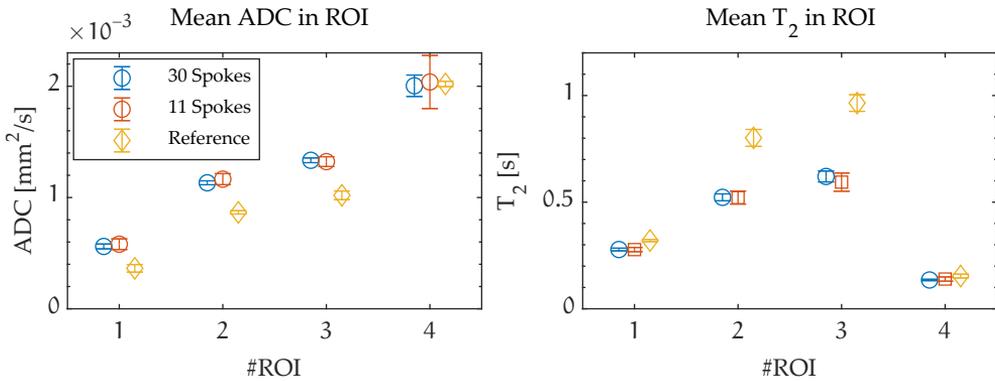


Figure 8.7: Region analysis of the data presented in figure 8.6. The position of the ROIs is indicated in figure 8.6. Systematic deviations from the reference measurements are seen for $ADCs < 2 \cdot 10^{-3} \text{ mm}^2/\text{s}$ and $T_2s > 200 \text{ ms}$. Figure taken from [BLB21, fig. 5].

echo, with 30 echoes and 30 spokes per echo. The full parameter setup is found in listing 11.4 in the appendix.

Using the optimized protocol, a measurement of the brain of a healthy volunteer was performed. Additionally, reference ADC (ssEPI) and T₂ (multiecho FSE) maps were acquired, as well as a high resolution T₂-weighted FSE for anatomical reference. For image reconstruction, only spokes with a $b < 800 \text{ s}/\text{mm}^2$ and $TE < 300 \text{ ms}$ were used, resulting in a total of 82 different b-TE-pairs. Other echoes were excluded because they were too low in signal magnitude and thus contributed noise only. The measurement was performed with a 20 channel head coil, of which 4 channels were excluded due to their minimal contribution of signal. The reconstruction used 10 iterations, and the interleaved model with a wavelet regularization and a regularization factor $\lambda = 10^{-3}$. The results are shown in figure 8.8. The T₂ map and the reference T₂ map are in good agreement in gray and white matter, with errors not exceeding 35 ms. Figure 8.9 shows the results from a region based analysis in gray- and white-matter. Therefore, portions of gray and white matter were manually segmented on the anatomical T₂-weighted high-resolution reference image, and subsequently copied to the SATM and reference T₂-maps. The histogram validates the agreement between the two sequences: The mean gray matter T₂ values were $91.9 \pm 12.9 \text{ ms}$ and $95.0 \pm 18.0 \text{ ms}$ for the SATM and reference sequence, respectively, and $88.8 \pm 9.4 \text{ ms}$ and $84.5 \pm 12.0 \text{ ms}$ for white matter.

In blood vessels, larger deviations can be seen in the difference image. There are two possible reasons for these deviations: First, there is a geometric shift between the two images, the reference image being shifted a few pixels upwards. Small blood vessels with a width of a few pixels cannot be represented correctly in the difference image. Secondly, the new T₂ map is blurred, rendering the

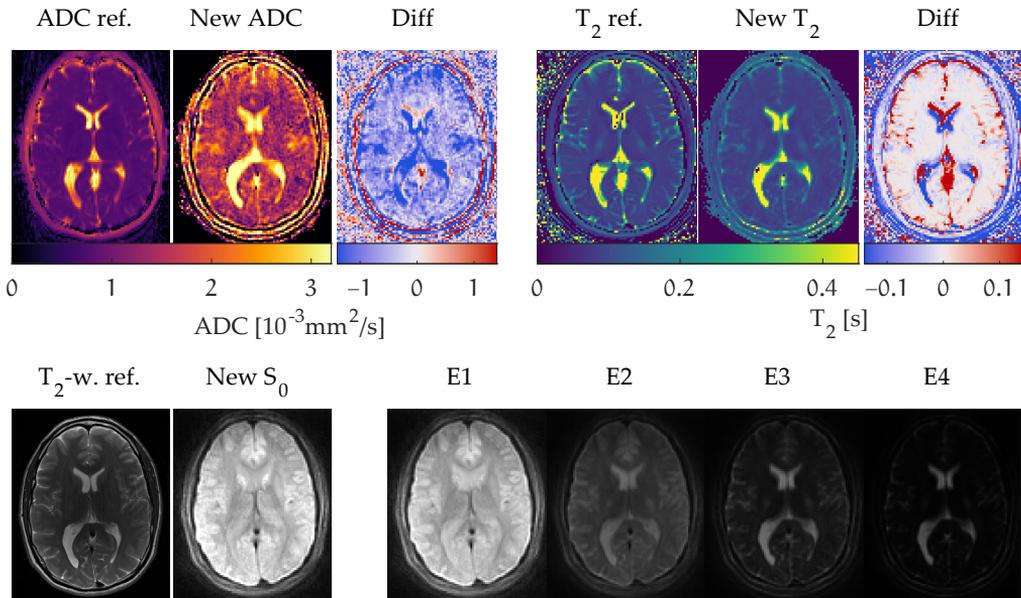


Figure 8.8: In vivo images of a healthy volunteer. The figure compares the parameter maps of the new sequence against conventional multiecho spin echo for T₂ mapping and DW-ssEPI for ADC mapping. Anatomical structure is compared against a T₂-w. reference image acquired with a high resolution FSE sequence. Four examples of the reconstruction with TE = [27, 110, 189, 297] ms of the in total 82 echoes are shown on the bottom right.

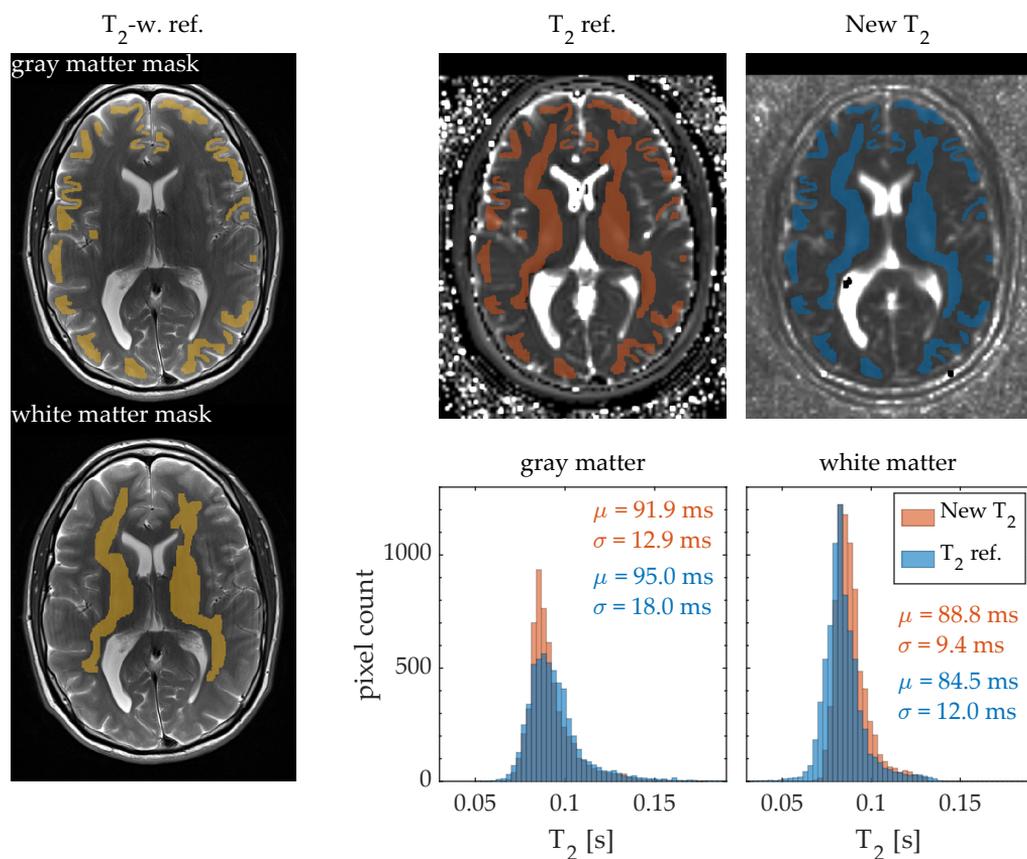


Figure 8.9: Region-based analysis in gray and white matter. The region masks were manually drawn on the high-resolution T₂-weighted FSE (left). The SATM and reference sequences were registered to the high-resolution images and the mask was subsequently used on the T₂-maps (top right). Based on the region masks, a histogram analysis was performed and mean and standard deviation values were extracted (bottom right).

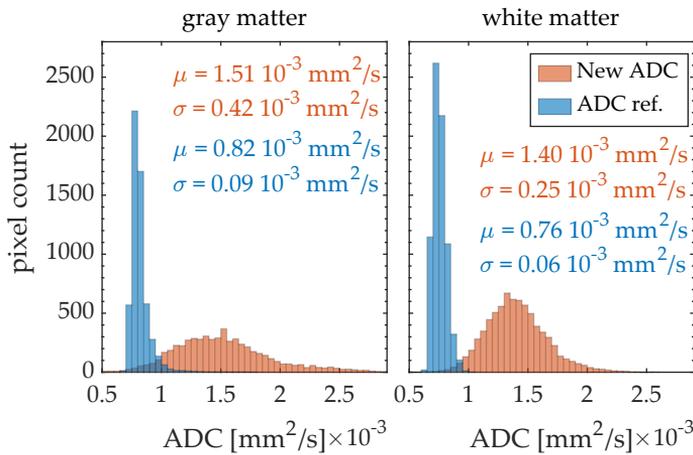


Figure 8.10: Region based analysis in gray and white matter for ADC maps. The region masks were taken from figure 8.9 and used on the ADC maps. The histogram analysis shows large deviations between SATM and reference ADC maps.

hyperintense fine structure less intensive and less sharp. For diffusion, the sequence generally overestimates the ADC values compared to the reference sequence. The region based analysis for the ADC maps was performed as described before for the T_2 maps, and the results are shown in figure 8.10. Mean values of $1.51 \pm 0.42 \cdot 10^{-3} \text{ mm}^2/\text{s}$ in gray matter were calculated for the SATM sequence, while the reference sequence yields $0.82 \pm 0.09 \cdot 10^{-3} \text{ mm}^2/\text{s}$. For white matter, the mean values are $1.40 \pm 0.25 \cdot 10^{-3} \text{ mm}^2/\text{s}$ and $0.76 \pm 0.06 \cdot 10^{-3} \text{ mm}^2/\text{s}$, respectively. Similar to the T_2 map, the blurring of the image leads to greater deviations in small blood vessels of up to $1.7 \cdot 10^{-3} \text{ mm}^2/\text{s}$. Geometric distortions can be seen in the reference ADC map, especially in the anterior part of the brain, which are not present in the SATM sequence (figure 8.8). This is confirmed by comparing the S_0 map to the high resolution T_2 -weighted image, since the new T_2 map, ADC map and S_0 map share the same geometry. In regions with very long T_2 , such as the CSF, the reference and the new method disagree both for T_2 and for ADC values.

Figure 8.8 also shows 4 synthetic echo images E1–E4, with $TE = [27, 110, 189, 297] \text{ ms}$ and $b = [66, 239, 197, 263] \cdot 10^{-3} \text{ s}/\text{mm}^2$, respectively. In total, there are 82 synthetic echo images, corresponding to each unique b - TE -pair. The contrast varies from predominantly proton density with weak T_2 and diffusion-weighting to a mixture of strong T_2 and diffusion-weighting.

Part IV

DISCUSSION AND OUTLOOK

DISCUSSION

9.1 AUTOMATED TUMOR SEGMENTATION

With the increasing usage of automated lesion detection and segmentation algorithms in clinical practice, the optimization of CNN architectures and training processes has been a focus topic in the last decade. However, there is only limited research on the optimization of the MR acquisition parameters with respect to automatic segmentation processes [Wah21; WW20].

In this thesis, CNNs were developed and trained to segment head and neck tumors in multiparametric MRI data. Therefore, MRI data was used from the prospective FMISO trial with up to 7 different image contrasts for CNN training. To increase image quality, as well as data quantity for CNN training, an improved patient setup was developed, which decreased Nyquist ghost artifacts in DW-MRI.

Using the CNNs, the influence of variations of the input MRI data on the segmentation performance was tested. To estimate the benefits of an additional distortion correction pre-processing step, the same CNNs were trained without and with distortion corrected ADC data (CNN-DistCorr). The second experiment, CNN-Info, characterized the relative information content of each input sequence in a 7-channel multiparametric MRI segmentation CNN.

The following sections are based on the works in [Bie19b] (section 9.1.2) and [Bie20b] (section 9.1.3).

9.1.1 B_0 FIELD HOMOGENIZATION

B_0 homogeneity in the head and neck region was improved by a modified patient setup. First, the setup proposed by Oudeman et. al. [Oud16] was reproduced, and B_0 field map measurements were performed on healthy volunteers. The setup significantly increased B_0 homogeneity in plane (axial direction) as well as in head-foot direction.

Due to the U-shaped pillow, which was placed on the anterior part of the neck, this setup had to be modified to account for the increased pressure and pain sensitivity of patients with head and neck cancer. The modified setup included two separate water bags on each side of the neck, with no additional load on the patients neck. It was evaluated using field map measurements in a volunteer, resulting in the expected increased B_0 homogeneity.

The new setup was integrated in the FMISO MRI protocol and a statistical analysis on the image quality, as assessed by the amount of Nyquist ghost artifacts, was performed. Images of ssEPI and rsEPI were graded according to the severity of the artifact, and the patient cohort with the new setup ($N = 22$) showed significantly better image quality than those without ($N = 54$). The combination of rsEPI and the improved setup yielded the best results, with 95% of the images graded as "good" as opposed to 74% before.

There is a wide range of other solutions to reduce B_0 inhomogeneity-related artifacts in the head and neck region. For example, the strong field variation in head-foot direction can be reduced by an adaptive slice-by-slice shimming [BRN96; MS97; Han17; Qiu21]. Slice-by-slice shimming can, however, lead to inconsistent slice positioning which is hard to recover without knowledge of the original field distribution and the shim. This inconsistency must be avoided as the radiation therapy planning is performed on the inconsistent images, which can potentially lead to excessive damage of nearby, healthy tissue. Since the global shim coils of the MRI system usually do not allow for more than second order corrections in transverse and fifth order in axial direction, local shimming was suggested as a means to counter the strong local field variations in the head and neck area [Pfa18; Gat16; Wal17]. The hardware necessary for these corrections has become available recently and is integrated into specialized head and neck coil arrays, which are incompatible with the thermoplastic fixation mask needed for patient immobilization. On the other hand, the setup using the water bags is easily implemented independent of vendor-specific sequences or additional shimming hardware.

9.1.2 DISTORTION CORRECTION

The CNN-DistCorr experiment compared tumor segmentations of a CNN trained on ADC images which were subject to different pre-processing steps. The network was trained on 5 MRI input channels and two different settings were compared: In the first case, the ADC data were subject to geometric distortion correction, and in the second case they were used without further processing. The CNNs were trained on data from 18 patients, but no significant

difference could be found between the networks trained on corrected and uncorrected data.

The distortion correction algorithm in this work is based on local mutual information registrations of the underlying DW-MRI and anatomically precise T_2 -weighted images. The algorithm severely reduced image distortion, as tested in simulations and experimental data that was validated by additional field map measurements. Still, the registration algorithm failed at locations of strong B_0 gradients, where pixels from multiple locations in the ADC images can be mapped to the same position. However, neither mutual information based, nor the field map based distortion correction technique could recover the undistorted image at such locations, which mostly appear at sharp air-tissue boundaries.

ADC images are known to be valuable inputs for segmentation of other tumor entities, like prostate or breast [ST16; Yabo8; Buro9], and similar findings were reported for head and neck cancer [Dri15; CS18]. Therefore, it is surprising that the distortion correction pre-processing step did not show clear improvement of the CNNs segmentation performance. However, there are several possible reasons for this result:

First, although the experimental data did not show a statistically significant difference in the two sets, the low number of available datasets ($N = 18$), might not be sufficient to reveal more subtle differences at all. The average segmentation performance, as measured by the DSC, was 0.40 ± 0.18 and 0.37 ± 0.21 for distortion corrected and uncorrected input data respectively. The high variation in the segmentation performance is an indication that a generally better performing network is necessary to thoroughly answer the question of the influence of distortion correction. This increase in performance may be achieved by increasing the amount of training data.

Secondly, the contribution of the other 4 input channels may mask the influence of geometric distortions in the ADC channel: With high-resolution anatomical data as seen in the T_1 and T_2 -weighted FSE, the DWI information might only add low-resolution features to the segmentation. A combination of these low-resolution ADC features, and high-resolution features from other input channels can lead to a good segmentation even though the ADC images are distorted. However, with the introduction of multiple contrasts, additional geometric inconsistencies arise. For example, when the patient swallows during the examination, tumor borders can shift with respect to the reference image, as shown in figure 9.1. Hence, even high-resolution image data may not be geometrically precise with respect to the tumor borders.

Lastly, the effect of distortion within the tumors is much smaller than in the rest of the image: Within the primary tumor, the distortion correction

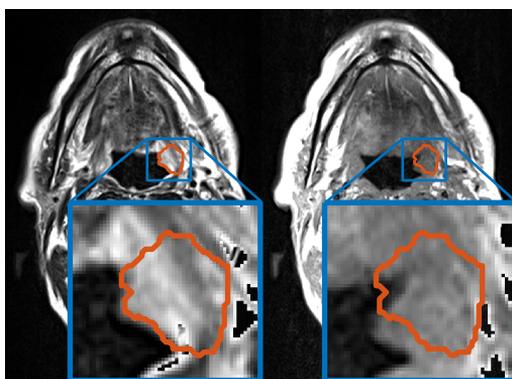


Figure 9.1: T_2 -weighted (left) and T_1 -weighted (right) image of a HNSCC patient. The tumor contour was drawn on the T_2 -weighted image. Because the patient moved during the examination (presumably swallowing), the tumor moved with respect to the contour.

maps calculated with the proposed algorithm yielded a standard deviation of 2.29 pixels, while it was 4.28 pixels for the complete image on average. Hence, strongly distorted areas tend to be outside the region of interest and might not contribute significantly to the segmentation performance.

The primary limitation of this study is the small amount of image data. Hence, to increase statistical significance, a 14-fold leave-one-out cross testing scheme was employed after the CNN architecture and hyperparameters were finalized. This testing scheme allowed a training set size of 13 patients and only a single test patient per network, and, therefore statistically relevant results could be drawn from the analysis. A major drawback of the cross testing scheme is the additional time needed for training: here, a single network trained for 50-75 h, and a total of 28 networks needed to be trained.

To further improve the networks performance, ADC maps should be acquired distortion-free already during image acquisition. A method to achieve geometrically more accurate results has been presented in the previous section. Additionally, increased SNR is beneficial for the segmentation performance [Bie20a], and can be achieved by additional signal averaging. However, this leads to longer acquisition times and cannot compensate the ADC bias due to magnitude signal averaging (see section 3.2.3). Alternatively, noise can be included in the model-fit routine, which can reduce the heterogeneity of the ADC images [JRS16; Walog]. Additional optimization of the b-values can then further improve the SNR in the target regions [SLN11; BB17].

9.1.3 CHANNEL-WISE IMPORTANCE ANALYSIS

In the CNN-DistCorr experiment, the influence of the different input channels on the segmentation performance was generally unclear, which introduced some ambiguity to the interpretation of the distortion correction. Therefore, in the CNN-Info experiment, the influence of different input channels with respect to the segmentation performance was analyzed for primary tumor and lymph node metastases.

To include as many MRI input channels as possible, v_e perfusion maps and T_1 -weighted post contrast DIXON water images from the FMISO trial were added as input to the network. To compensate for the small size of the dataset, additional patient data during radiation treatment was used as independent datasets. Moreover, 6 additional patients were included in the trial that were used for the analysis, resulting in a total of 36 datasets from 18 different patients. 13 of the datasets were acquired before treatment start (week 0), 9 were acquired at week 2 and 14 at week 5 during treatment.

Next, 8 different CNNs were trained: one CNN included all 7 input channels, and the other CNNs were trained each with one of the channels left out. Then, the segmentation performance of the LOO-CNNs was compared to the reference CNN to estimate the influence on segmentation performance for each of the channels left out.

The overall segmentation performance was similar to CNN-DistCorr, with best scores of 65% for primary tumor, and 58% for lymph node metastases. On average, the networks showed segmentation performances of 30% and 24%, respectively. The reduced average performance as compared to the first experiment can be explained by the changes made in the input dataset: in CNN-Info, tumors were included which can express systematic changes of their contrast parameters as an effect of the radiochemotherapy [Wie19; Wie20]. Additionally, the segmentation performance correlates with tumor size, and, thus, the inclusion of tumors that have shrunk during treatment is expected to decrease the average segmentation performance. This effect may be compensated by using a mean border distance metric in addition to the DSC, e. g., the Average Hausdorff Distance or the Probabilistic Distance [GJC01; TH15].

Furthermore, the inclusion of additional data increases the statistical validity of the study, as deviations from a common mean value are investigated, and not the absolute segmentation performance itself. Thus, the difference in segmentation performance was computed for pairs of LOO-CNNs and reference CNNs with identical input patient sets. The highest statistically significant influence was found in the T_2^* -channel, with an average decrease

of $\Delta\text{DSC}_{\text{GTV-T}} = 5.7\%$ and $\Delta\text{DSC}_{\text{GTV-Ln}} = 5.8\%$ for primary tumor and lymph node metastases respectively. This was followed by the T_1 -weighted contrast enhanced ($\Delta\text{DSC}_{\text{GTV-T}} = 5.0\%$) and v_e ($\Delta\text{DSC}_{\text{GTV-T}} = 4.9\%$) channels for tumor segmentation. Although the other results did not reach statistical significance ($p < 0.05$ in the paired Students t-test), the T_1 -weighted pre-contrast channel showed a $\Delta\text{DSC}_{\text{GTV-Ln}} = 4.9\%$ for the segmentation of lymph node metastases, scoring second after T_2^* . Surprisingly, the ADC channel showed the least overall influence on the segmentation performance ($\Delta\text{DSC}_{\text{GTV-T}} = 2.4\%$, $\Delta\text{DSC}_{\text{GTV-Ln}} = 2.2\%$), however without statistical significance. This may be explained by the higher number of small tumors included in the study in combination with the geometric distortion artifacts in ADC maps: The relative effect of the distortion is larger in smaller structures, which can lead to a decreased significance in the CNNs decision-making process. On the other hand, this result validates the observations made in the CNN-DistCorr experiments, where changes in the ADC dataset did not lead to a significant change in the segmentation performance.

This experiment for the first time compared the influence of 7 different MRI channels to the segmentation performance of a CNN. The results allow an optimization of the measurement protocol for head and neck cancer patients for automated tumor segmentation of the images.

The results shown in this work are promising, yet other groups reach segmentation performances of beyond 90% [Lit17b; Akk17; Sha20] in other body regions, or focus on the segmentation of organs at risk in CT images [MDL17; van19]. In the head and neck area, only few studies performed automatic tumor segmentation, with a DSC scores of up to 0.73 based on only T_1 -weighted and T_2 -weighted images [Wah22], or 0.75 based on PET/CT images [Moe19]. However, Wahid et. al. [Wah22] selected from a dataset of 124 HNSCC patients a subgroup of 30 oropharyngeal cancer patients, a distinction which was not made in the present dataset.

There are several reasons for the discrepancy in performance to the results published in the literature: Often, good segmentation results are achieved in tumor entities with clearly differentiated contrast to the surrounding tissue, such as brain tumor [Kam17]. Many of those groups use large public databases to train high-performing segmentation networks. Here, no such database exists, because MRI with more than T_1 -weighted, T_2 -weighted and DWI are seldomly performed in clinical practice. Thus, data needs to be taken from smaller patient trials with highly customized imaging protocols, which can be inefficient with respect to imaging time for the patient treatment. In contrast to head and neck cancer, brain tumors can be considered a true rigid body problem, which makes co-registration of different input channels accurate and easy to implement. The

head and neck area, on the other hand, is subject to local deformations, e. g. by tongue movement or swallowing motion, which results in inaccuracies in image registration. In this study, additional patient motion must be expected due to the long acquisition times of about 40 minutes, adding to the limitation of co-registration.

Although the limitations in co-registration appear very severe, the same motion and co-registration challenges have to be expected between MRI and radiation therapy, which sets an upper limit to the necessary accuracy of dose application [Yan98]. Hence, a segmentation performance that surpasses this intrinsic limitation is not expected to increase treatment outcome.

9.1.4 CONCLUSION

In summary, it was shown that within the anatomically highly challenging head and neck region, multiparametric MRI can be used to achieve good automatic tumor segmentations with CNNs. The influence of different variations in the input data was analyzed, and a non-significant decrease in segmentation performance was observed for geometrically uncorrected input diffusion data. Moreover, if any of the 7 possible input channels were left out, the segmentation performance decreased as well.

Tools for the fast and accurate automatic contouring of tumors in the head and neck area will become indispensable in future application: MR-guided radiotherapy based on MR-LINAC systems promises better tumor control by real time replanning of dose delivery [Chu18]. Such concepts automatically take into account the strong changes in the patients anatomy during radiochemotherapy, thereby minimizing the risk for nearby organs [Sur17; Kat].

Daily or real-time imaging-adapted treatment plans can not be provided by the tedious manual segmentation processes, which is the current state of the art. Hence, this thesis helps in the successful combination and implementation of MRI protocols optimized for CNN based, automatic tumor segmentation.

9.2 SIMULTANEOUS ADC AND T₂ MAPPING

A sequence for simultaneous acquisition of T₂ and ADC parameter maps (SATM) was developed and tested. The sequence was designed to acquire T₂ as another potentially highly sensitive parameter for automatic tumor segmentation in head and neck cancer, while removing some of the inherent artifacts in DW-EPI imaging for ADC quantification. Primary design goals were acquisition speed and geometric accuracy, next to precise T₂ and ADC quantification.

The simulations of the sequence show that images can be reconstructed based on only 11 spokes per b-TE-pair and 45 b-TE-pairs, such that a volume of 16 slices could be acquired in less than 3 minutes, using a TR of 5 s.

A major problem in multiecho SE sequences is the interference of different echo pathways. Using a phase graph analysis, a soft CPMG-condition was derived to ensure coinciding echoes despite the violation of the second CPMG condition. Bloch simulations showed that the artifact caused by the CPMG violation could be avoided, recovering up to 5 % of signal variation. However, random patient motion remains a challenge for both, conventional multiecho SE and the SATM sequence, as moving spins violate the CPMG conditions either way.

For reconstruction, two different methods were developed and compared. Both methods are based on an iterative minimization of a loss function. The first method directly models the measurement as a function of the target parameters ADC, T_2 and S_0 . The second method interleaves a model-free image intensity based reconstruction, with an image update by fitting the bi-exponential ADC- T_2 model to the current data. The second reconstruction method (interleaved model) proved to be faster and more robust in simulations as well as on experimental data. The first reconstruction method (iterative model) has a poorer performance because it models each image phase as a free model parameter. This limitation is necessary due to the random phase artifact, which can appear in DWI. Thus, the number of free model parameters is largely increased as compared to the interleaved model. Still, for both models k-space spokes from different excitations could be combined because the phase artifacts were less severe than in conventional DWI. This might be due to the small increase in b-value of 80-150 s/mm² per diffusion block in the SATM sequence, as opposed to a single block with a b-value of up to 800 s/mm² in conventional DWI.

In simulations, both reconstruction methods failed to reconstruct correct ADC values in short T_2 regimes ($T_2 < 100$ ms), which could be caused by the strong signal attenuation due to T_2 -relaxation. In this sequence, ADC is inherently more difficult to estimate than T_2 due to the lower amount of unique b-values compared to unique TEs. Thus, the errors on the ADC estimation are expected to be higher than those for T_2 estimation.

Phantom measurements showed that the T_2 fit additionally shows a strong systematic error of up to 35% for reference values $T_2 > 800$ ms. However, such long relaxation times are generally not encountered in head and neck tumor imaging. Consequently, the SATM sequence was not optimized for these values, reaching maximum TEs of 300 ms.

The sequence was then optimized for *in vivo* imaging, with 82 unique b-TE-pairs with TEs between 27-300 ms and b-values between 68-800 mm²/s. In gray and white matter, a good agreement with the reference measurements was observed, while for long T_2 systematic differences were found. However, this error in T_2 quantification was expected from the previous simulations and phantom measurements.

Similarly, the proposed ADC measurements show a systematic overestimation compared to the reference. This deviation may be caused by simultaneously occurring (stimulated) echoes, which experienced different signal attenuation, but are not modeled in the reconstruction. A good slice profile reduces this deviation, but assuming a certain residual B_1 inhomogeneity and an imperfect slice profile, signal contributions from non-180° RF-pulses cannot be avoided.

The parameter maps show excellent anatomical agreement to the reference T_2 -weighted FSE sequence, validating the geometrical accuracy of the new ADC map.

Simultaneous T_2 and ADC imaging has been investigated in other studies before: [Sta12] and [Gra17] used a DESS-based sequence with additional diffusion weighting. However, the DESS sequences showed a relatively long acquisition time of more than 10 minutes. [Ali18] proposed to simultaneously measure T_2 and ADC using a diffusion weighted spin-echo EPI based sequence. Similarly, [SBE17] proposed a EPI-readout based sequence for simultaneous T_2 and ADC estimation. Their methods, however, suffer from the same geometric distortion artifacts as conventional DWI, and therefore no advantages for head and neck cancer imaging are expected. The SATM sequence, on the other hand, solved the geometrical inaccuracies and can acquire T_2 and ADC maps in a shorter amount of time.

OUTLOOK

In this thesis, the influence of MRI input channels on the segmentation performance of a CNN was analyzed: First, with regard to the geometric accuracy of the input ADC-maps, and, second, with regard to the relative information content of 7 different channels.

In future studies, this relative importance analysis can be extended: The overall segmentation performance has to be increased by increasing the size of the dataset on the one hand, and by comparing different network architectures (e. g., *DeepMedic* and *U-Net*) on the other hand. With ongoing patient recruitment in the FMISO trial, the existing database can be increased up to 70 patients in total. Then, the gradient-weighted class activation mapping (Grad-CAM) algorithm [Sel17] should be adapted for channel-wise activation maps, which could additionally quantify the relative importance of each MR contrast in a single CNN simultaneously. Next, the effects of different input combinations should be tested on CNNs with reduced numbers of channels, e. g., a CNN configuration without any anatomical or without any functional input data. Finally, the resulting protocol optimization has to be tested in a realistic setting: Therefore, patient data should be acquired with an improved protocol, and the segmentation performance has to be tested against the previous configuration.

Additionally, the SATM sequence can be used to acquire additional T_2 and ADC parameter maps. Therefore, the sequence has to be optimized to remove any systematic bias in ADC and T_2 quantification. Hence, a thorough theoretical description of the signal attenuation due to diffusion and T_2 -decay has to be formulated, similar to [WB90]. Furthermore, the single radial readout spokes could be replaced by a spiral trajectory, transforming the sequence into a single shot sequence. With these improvements, the SATM sequence may potentially replace the conventional DW-MRI as well as the T_2 -weighted MRI in the FMISO protocol with reliable, fast and distortion-free T_2 and ADC maps, or, if advantageous for the CNN segmentation, synthetic T_2 and ADC weighted images. With reliable parameter mapping, the sequence may also prove to be beneficial in other tumor entities.

Part V

APPENDICES

SUPPLEMENTARY MATERIAL

11.1 AUTOMATED TUMOR SEGMENTATION

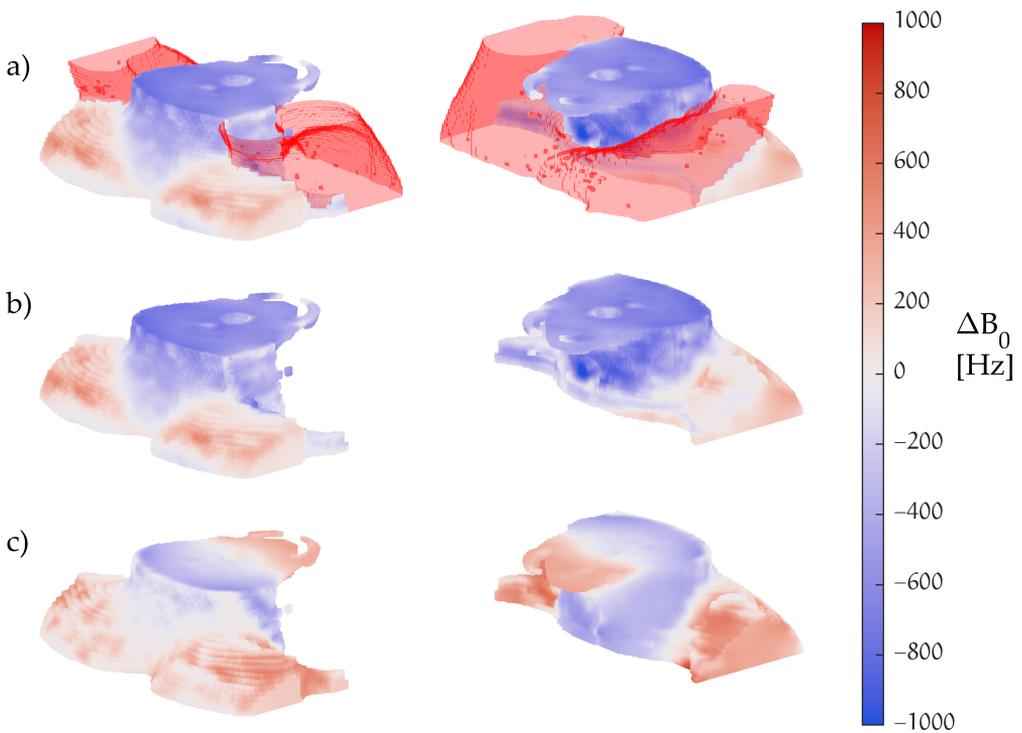


Figure 11.1: Fieldmap measurements in a second volunteer, using a U-shaped neck pillow to increase the B_0 homogeneity. a) shows the field map with the neck pillow, while b) shows the same data with the pillow removed for visibility. c) shows the measurement without the neck pillow, with significantly increase B_0 inhomogeneity.

Listing 11.1: CNN model configuration file used for the CNN-DistCorr experiment.

```
#Default values are set internally, if the corresponding parameter is not
  found in the configuration file.

#[Optional but highly suggested] The name will be used in the filenames
  when saving the model.
#Default: "cnnModel"
modelName = "deepMedicTumorMasked_o8"

#[Required] The main folder that the output will be placed.
folderForOutput = "/bielak/fmiso/deepMedicTumorMasked_o8_output"

#===== MODEL PARAMETERS =====

#[Required] The number of classes in the task. Including background!
numberOfOutputClasses = 2
#[Required] The number of input channels, eg number of MRI modalities.
numberOfInputChannels = 5

#++++++Normal pathway+++++
#[Required] This list should have as many entries as the number of layers
  I want the normal-pathway to have.
#Each entry is an integer that specifies the number of Feature Maps to
  use in each of the layers.
numberFMsPerLayerNormal = [30, 30, 40, 40, 50, 50, 70, 70]
#[Required] This list should have as many entries as the number of layers
  in the normal pathway.
#Each entry should be a sublist with 3 entries. These should specify the
  dimensions of the kernel at the corresponding layer.
kernelDimPerLayerNormal = [[3,3,3], [3,3,3], [3,3,3], [3,3,3], [3,3,3],
  [3,3,3], [3,3,3], [3,3,3]]

#[Optional] List with number of layers, at the output of which to make a
  residual connection with the input of the previous layer. Ala Kaiming
  He et al, "Deep Residual Learning for Image Recognition".
#Note: Numbering starts from 1 for the first layer, which is not an
  acceptable value (no previous layer).
#Example: [4,6,8] will connect (add) to the output of Layer 4 the input
  of Layer 3. Also, input to 5th will be added to output of 6th, and
  input of 7th to output of 8th.
#Default: [], no residual connections
layersWithResidualConnNormal = []
```

```

#[Optional] Layers to make of lower rank. Ala Yani Ioannou et al,
    "Training CNNs with Low-Rank Filters For Efficient Image
    Classification".
#Example: [3,5] will make the 3rd and 5th layers of lower rank.
#Default: []
lowerRankLayersNormal = []

#++++++Subsampled pathway+++++
#[Optional] Specify whether to use a subsampled pathway. If False, all
    subsampled-related parameters will be read but disregarded in the
    model-construction.
#Default: False
useSubsampledPathway = True

#[Optionals] The below parameters specify the subsampled-pathway
    architecture in a similar way as the normal.
#If they are omitted and useSubsampledPathway is set to True, the
    subsampled pathway will be made similar to the normal pathway
    (suggested for easy use).
#[WARN] Subsampled pathway MUST have the same size of receptive field as
    the normal. Limitation in the code. User could easily specify
    different number of FMs. But care must be given if number of layers
    is changed. In this case, kernel sizes should also be adjusted to
    achieve same size of Rec.Field.
numberFMsPerLayerSubsampled = [30, 30, 40, 40, 50, 50, 70, 70]
kernelDimPerLayerSubsampled = [[3,3,3], [3,3,3], [3,3,3], [3,3,3],
    [3,3,3], [3,3,3], [3,3,3], [3,3,3]]

#[Optional] How much to downsample the image that the subsampled-pathway
    processes.
#Default: [3,3,3]
subsampleFactor = [3,3,3]

#[Optional] Residual Connections for subsampled pathway. See
    corresponding parameter for normal pathway.
#Default: mirrors the normal pathway, no residual connections
#layersWithResidualConnSubsampled = []

#[Optional] Layers to make of lower rank. See corresponding parameter for
    normal pathway.
#Default: Mirrors the normal pathway
#lowerRankLayersSubsampled = []

#++++++FC Layers+++++

```

```

#[Optional] After the last layers of the normal and subsampled pathways
    are concatenated, additional Fully Connected hidden layers can be
    added before the final classification layer.
#Specify a list, with as many entries as the number of ADDITIONAL FC
    layers (other than the classification layer) to add. The entries
    specify the number of Feature Maps to use.
#Default: []
numberFMSPerLayerFC = [100, 100]

#[Optional] Specify dimensions of the kernel in the first FC layer. This
    kernel combines the features from multiple scales. Applies to the
    final Classification layer if no hidden FC layers in network.
#Note: convolution with this kernel retains the size of the FMs (input is
    padded).
#Default: [1,1,1]
kernelDimFor1stFcLayer = [1,1,1]

#[Optional] Residual Connections for the FC hidden layers. See
    corresponding parameter for normal pathway.
#Default: [], no connections.
#layersWithResidualConnFC = []

#+++++++Size of Image Segments+++++++
#DeepMedic does not process patches of the image, but larger
    image-segments. Specify their size here.

#[Required] Size of training segments influence the captured distribution
    of samples from the different classes (see DeepMedic paper)
segmentsDimTrain = [20,20,20]
#[Optional] The size of segments to use during the validation-on-samples
    process that is performed throughout training if requested.
#Default: equal to receptive field, to validate on patches.
segmentsDimVal = [17,17,17]
#[Optional] Bigger image segments for Inference are safe to use and only
    speed up the process. Only limitation is the GPU memory.
#Default: equal to the training segment.
segmentsDimInference = [45,45,45]

#+++++++Batch Sizes+++++++
#[Required] The number of segments to create a batch.
#The samples in a training-batch are all processed and one optimization
    step is performed.
#Larger batches approximate the total data better and should positively
    impact optimization but are computationally more expensive (time and
    memory).
batchSizeTrain = 10

```

```
#[Optionals] Batch sizes for validation and inference only influence the
    speed. The bigger the better. Depends on the segment size and the
    model size how big batches can be fit in memory.
#Default: Equal to train-batch size.
batchSizeVal = 48
batchSizeInfer = 10

#[Optionals] Dropout Rates on the input connections of the various
    layers. Each list should have as many entries as the number of layers
    in the corresponding pathway.
# 0 = no dropout. 1= 100% drop of the neurons. Empty list for no dropout.
#Default: []
dropoutRatesNormal = []
dropoutRatesSubsampled = []
#Default: 50% dropout on every Fully Connected layer except for the first
    one after the concatenation
#Note: The list for FC rates should have one additional entry in
    comparison to "numberFMsPerLayerFC", for the classification layer.
dropoutRatesFc = [0.0, 0.5, 0.5] # +1 for the classification layer!

#[Optional] Initialization method of the kernel weights. Specify 0 for
    classic, from the normal distribution  $N(0, 0.01)$ . Otherwise specify 1
    for the method of He et al from "Delving Deep into Rectifiers".
#Default: 1
initializeClassic0orDelving1 = 1
#[Optional] Activation Function for all convolutional layers. Specify 0
    for ReLU, 1 for PreLU.
#Default: 1
relu0orPrelu1 = 1

#[Optional] Batch Normalization uses a rolling average of the mus and std
    for inference. Specify over how many batches (optimization steps)
    this rolling average should be taken.
#Default : 60 (in our usual settings, with batchsize=10, segments per
    training subepoch=1000, and subepochs per epoch=20, this averages
    over 5 epochs).
rollAverageForBNOverThatManyBatches = 60
```

Listing 11.2: CNN training configuration file as used for the CNN-DistCorr experiment.

```

#Default values are set internally, if the corresponding parameter is not
  found in the configuration file.

#[Optional but highly suggested] The name will be used for saving the
  models,logs and results.
#Default: "trainSession"
sessionName = "trainSessionDeepMedic_01_Statistic_Corr"

#[Required] The main folder that the output will be placed.
folderForOutput = "/bielak/fmiso/DeepMedic_01_Statistic_Corr"

#[Optional] The path to the saved CNN-model to use for training. Optional
  in the case the the model is specified from command line with the
  -model option. In this case, this entry file of the config file will
  be disregarded, and the one from the command line will be used.
cnnModelFilePath =
  "/deepMedicTumorMasked_08.initial.2017-10-13.10.37.37.935859.save"

#=====Training=====

#++++++++Input+++++++

#[Required] A list that should contain as many entries as the channels of
  the input image (eg multi-modal MRI). The entries should be paths to
  files. Those files should be listing the paths to the corresponding
  channels for each training-case. (see example files).
channelsTraining = ["/01_trainADCMapReg.cfg", "/01_trainKtransReg.cfg",
  "/01_trainT1TseReg.cfg", "/01_trainT2starReg.cfg",
  "/01_trainT2TseReg.cfg"]

#[Required] The path to a file which should list paths to the Ground
  Truth labels of each training case.
gtLabelsTraining = "/01_trainGtLabels.cfg"

#++++++++Sampling+++++++

#[Optional] The path to a file, which should list paths to the
  Region-Of-Interest masks for each training case.
#If ROI masks are provided, under default-sampling settings, the training
  samples will be extracted only within it. Otherwise from whole volume.
#This mask is also used for calculating mu and std intensities for
  intensity-augmentation, if performed.
#roiMasksTraining = "/trainRoiMasks.cfg"

```

```

#####Advanced Sampling#####
#Note: Given variables in this "Advanced Sampling" section are
      disregarded if default settings are used, unless one sets:
      useDefaultTrainingSamplingFromGtAndRoi = False.

#[Optional] True in order to use default sampling for training. In this
      case, foreground samples are extracted from within the GT mask.
#Background samples are then extracted from the ROI (or full volume),
      excluding the GT. By default, equal number of samples are extracted
      from foreground and background.
#Note: Advanced options are disabled if default settings are used.
#Default: True
useDefaultTrainingSamplingFromGtAndRoi = True

#[Optional] Type-of-Sampling to use for training.
#[Possible Values] 0 = Foreground / Background, 1 = Uniform, 2 = Full
      Image , 3 = Separately-Per-Class.
#Note: In case of (2) Full Image, ensure you provide segmentsDimTrain in
      modelConfig.cfg at least as big as image dimensions (+CNN's receptive
      field if padding is used).
#Default: 0
typeOfSamplingForTraining = 0

#[Optional] List the proportion (0.0 to 1.0) of samples to extract from
      each category of samples.
#Note: Depending on the Type-of-Sampling chosen, list must be of the form:
#      >> Fore/Background: [proportion-of-FOREground-samples,
      proportion-of-BACKground-samples], eg [0.3, 0.7]. IMPORTANT:
      FOREground first, background second!
#      >> Uniform or Full-Image: Not Applicable and disregarded if given.
#      >> Separate sampling of each class:
      [proportion-of-class-0(background), ..., proportion-of-class-N]
#Note: Values will be internally normalized (to add up to 1.0).
#Default: Foreground/Background or Separately-Each-Class : equal number
      of segments extracted for each of the categories. Uniform or
      Full-Image: N/A
proportionOfSamplesToExtractPerCategoryTraining = [0.5, 0.5]

#[Optional] This variable allows providing weighted-maps to indicate
      where to extract more segments for each category of samples. Higher
      weight means more samples from that area.
#The value provided should be a List with paths to files. As many files
      as the categories of samples for the chosen Sampling-Type.
#Similarly to the files listing the Ground Truth, Channels, etc per
      subject, these files should list the paths to the weight-maps of each
      subject for the corresponding category.

```

```

#Note: Number of files required: Fore/Backgr:2, Uniform:1,
      Full-Image:N/A, Separate each class:NumOfOutputClasses (Incl Backgr).
#IMPORTANT: Sequence of weight-maps is important!
#>> If Fore/Background type of sampling, provide for the FOREground first!
#>> If Separately sampling each class, provide weightmap-files in the
      same sequence as the class-labels in your Ground Truth! Eg
      background-0 first, class-1 second, etc.
#Default : If this variable is not provided, samples are extracted based
      on the Ground-Truth labels and the ROI.
weightedMapsForSamplingEachCategoryTrain =
    ["/01_weightMapsForeground.cfg", "/01_weightMapsBackground.cfg"]

#++++++++Training Cycle (see documentation)+++++++

#[Optionals but highly suggested as they are model dependent.]
#How many epochs to train for. Default: 35
numberOfEpochs = 30
#How many subepochs comprise an epoch. Every subepoch I get Accuracy
      reported. Default: 20
numberOfSubepochs = 20
#Every subepoch, load the images from that many cases and extract new
      training samples. Default: 50
numOfCasesLoadedPerSubepoch = 13#Every subepoch, extract in total this
      many segments and load them on the GPU. Memory Limited. Default:
      1000
#Note: This number in combination with the batchSizeTraining, define the
      number of optimization steps per subepoch (=NumOfSegmentsOnGpu /
      BatchSize).
numberTrainingSegmentsLoadedOnGpuPerSubep = 1700
#++++++++Learning Rate Schedule+++++++

#[Optional] The type of schedule to use for Learning Rate annealing.
#0=Stable Decrease. 1=Auto (Lower LR when validation accuracy plateaus.
      Requires validation-on-samples). 2=Lower at predefined epochs.
      3=Exponentially decrease LR, linearly increase Mom.
#Note: Training Schedule is very important. We suggest running stable and
      observing training error, then lower LR when it plateaus. Otherwise,
      use exponential but make sure to train for enough epochs.
stable0orAuto1orPredefined2orExponential3LrSchedule = 2

#[For Stable + Auto + Predefined] By how much to divide LR when lowering.
      Default: 2
whenDecreasingDivideLrBy = 2.0

```

```

#[For Stable + Auto] How many epochs to wait before decreasing again. Set
    Zero to never lower LR. Default: 3
numEpochsToWaitBeforeLoweringLr = 3
#[For Auto] If validation accuracy increases more than this much, reset
    the waiting counter. Default: 0.0005
minIncreaseInValidationAccuracyThatResetsWaiting = 0.0005

#[Required for Predefined] At which epochs to lower LR.
predefinedSchedule = [10, 14, 17, 20, 23, 25, 29, 31.34, 37, 40, 43, 46]

#[Required for Exponential] [First epoch to start lowering from, value
    for LR to reach at last epoch, value for Mom to reach at last epoch]
exponentialSchedForLrAndMom = [12, 1.0/(2**(7)), 0.9]

#++++++Data Augmentation+++++

#[Optional] Specify whether to reflect the images by 50% probability in
    respect to the X/Y/Z axis. Default: [False, False, False]
reflectImagesPerAxis = [True,True,True]

#[Optional] Augmentation by changing the mean and std of training
    samples. Default: False
performIntAugm = False
#I' = (I + shift) * multi
#[Optionals] We sample the "shift" and "multi" variable for each sample
    from a Gaussian distribution. Specify the mu and std.
#Defaults : [0, 0.1] and [1.,0.]
sampleIntAugmShiftWithMuAndStd = [0, 0.1]
sampleIntAugmMultiWithMuAndStd = [1., 0.0]

#++++++Optimization+++++

#[Optionals]
#Initial Learning Rate. Default: 0.001.
learningRate = 0.001
#Optimizer to use. 0 for classic SGD, 1 for Adam, 2 for RmsProp. Default:
    2
sgd0orAdam1orRms2 = 2
#Type of momentum to use. 0 for standard momentum, 1 for Nesterov.
    Default: 1
classicMom0orNesterov1 = 1
#Momentum Value to use. Default: 0.6
momentumValue = 0.6
#Non-Normalized (0) or Normalized momentum (1). Bear in mind that
    Normalized mom may result in smaller gradients and might need
    relatively higher Learning Rate. Default: 1
momNonNorm0orNormalized1 = 1

```

```

#Parameters for RmsProp. Default: rho=0.9, e=10**(-4) (1e-6 blew up the
  gradients. Haven't tried 1e-5 yet).
rhoRms = 0.9
epsilonRms = 10**(-4)

#[Optionals] Regularization L1 and L2.
#Defaults: L1_reg = 0.000001, L2_reg = 0.0001
L1_reg = 0.0000005
L2_reg = 0.00005

#++++++Freeze Layers++++++
#[Optional] Specify layers the weights of which you wish to be kept fixed
  during training (eg to use weights from pre-training). First layer is
  1.
# One list for each of the normal, subsampled, and fully-connected (as
  1x1 convs) pathways. For instance, provide [1,2,3] to keep first 3
  layers fixed. [] or comment entry out to train all layers.
# Defaults: [] for the Normal and FC pathway. For the Subsampled pathway,
  if entry is not specified, we mirror the option used for the Normal
  pathway.
layersToFreezeNormal = []
layersToFreezeSubsampled = []
layersToFreezeFC = []

#====Validation=====

#[Optionals] Specify whether to perform validation on samples and
  full-inference every few epochs. Default: False for both.
performValidationOnSamplesThroughoutTraining = True
performFullInferenceOnValidationImagesEveryFewEpochs = True

#[Required] Similar to corresponding parameter for training, but points
  to cases for validation.
channelsValidation = ["/validation/o1_validationADCMMapReg.cfg",
  "/validation/o1_validationKtransReg.cfg",
  "/validation/o1_validationt1TseReg.cfg",
  "/validation/o1_validationT2starReg.cfg",
  "/validation/o1_validationt2TseReg.cfg"]

#[Required for validation on samples, optional for full-inference]
  Similar to corresponding parameter for training, but points to cases
  for validation.
gtLabelsValidation = "/validation/o1_validationGtLabels.cfg"

```

```
#[Required] Similar to corresponding parameter for training. Only influences how accurately the validation samples will represent whole data. Memory bounded.
```

```
#Default: 3000
```

```
numberValidationSegmentsLoadedOnGpuPerSubep = 5000
```

```
#[Optional] Similar to corresponding parameter for training
```

```
#roiMasksValidation = "./validation/validationRoiMasks.cfg"
```

```
#+++++Advanced Validation Sampling+++++
```

```
#Note: Given variables in this "Advanced Validation Sampling" section are disregarded if default settings are used, unless one sets: useDefaultUniformValidationSampling = False.
```

```
#[Optional] True in order to use default sampling for validation. Default is uniform sampling within the ROI (or whole volume if not provided).
```

```
#Note: Advanced options are disabled if default settings are used.
```

```
#Default: True
```

```
useDefaultUniformValidationSampling = True
```

```
#[Optional] Type-of-Sampling to use for Validation. See description of corresponding variable for training.
```

```
#Default: 1 (uniform sampling)
```

```
typeOfSamplingForVal = 1
```

```
#[Optional] List the proportion (0.0 to 1.0) of samples to extract from each category of samples. See description of corresponding variable for training.
```

```
#Default: Foreground/Background or Separately-Each-Class : equal number of segments extracted for each of the categories. Uniform or Full-Image: N/A
```

```
#proportionOfSamplesToExtractPerCategoryVal = [0.3, 0.7]
```

```
#[Optional]
```

```
#The following variable allows providing weighted-maps that indicate where to acquire more samples for each category/class. See description of corresponding variable for training.
```

```
#Default : If this variable is not provided, samples are extracted based on the Ground-Truth labels and the ROI.
```

```
weightedMapsForSamplingEachCategoryVal =
```

```
["./validation/o1_weightMapsForeground.cfg",  
"./validation/o1_weightMapsBackground.cfg"]
```

```
#+++++Full-Inference on validation cases+++++
```

```
#[Optional] How often (epochs) to perform full inference. It is time
    consuming... Default: 1
numberOfEpochsBetweenFullInferenceOnValImages = 5

#[Optionals] Specify whether to save the segmentation and probability
    maps for each class. Default: True to all
saveSegmentationVal = True
saveProbMapsForEachClassVal = [True, True, True, True, True]

#[Required if requested to save results] The path to a file, which should
    list names for each validation case, to name the results after.
namesForPredictionsPerCaseVal =
    "./validation/01_validationNamesOfPredictions.cfg"

#--Feature Maps--
#Feature maps can also be saved, but section is omitted here. See testing
    configuration.

#====Generic====
#[Optional] Pad images to fully convolve. Default: True
padInputImagesBool = True
```

11.2 SIMULTANEOUS ADC AND T₂ MAPPING

Listing 11.3: MATLAB Code for the Bloch simulation. Credits: Dr. Jochen Leupold (original function for a RARE sequence), Lars Bielak (adaptation for SATM sequence).

```
function [Echoes]=BlochSim(nX,nEcho,nDiff,alpha,T1,T2,ES,diffgrad_factor)

%parameters:
% no_pulses - echo train length, corresponds to number of samples in both
    R0 and PE direction
% alpha - refocusing flip angle (in deg)
% T1
% T2
% ES - echo spacing
% diffgrad_factor - zeroth moment of one diffusion gradient relative to
    crusher gradient

linekpace=zeros(nX,nX,nEcho,nX);
phant=phantom(nX,nX); %make phantom
start=180; %introduces symmetry to dephasing (see dephase_function)
offres=0; %makes an off-resonance angle - recommendation leave it at zero
readgrad = 0.5; %moment of the readout prephaser after excitation pulse
crushergrad = 4*readgrad; %moment of the crusher gradients before and
    after each pulse
diffgrad = diffgrad_factor * crushergrad; %factor compared to readoutgrad
phasecyc = 180*ones(1,nEcho+1);phasecyc(1)=90; %cpmg phase cycle
no_isos=360*nX; %total number of magnetization vectors (360 per voxel)
echosamples=nX; % number of sampling points in R0 direction

for rep=1:nX % repetitions loop. The experiment is repeated for each
    phase encoding step
    for li=1:nX %line loop %loops over the lines in image space. idea: a
        kspace is constructed for any line, all these k-spaces ar added
        afterwards, finally building the acquired k-space of the phantom
        M_mat=zeros(no_isos,2); % Matrix containing magnetisation.
            (develops through the whole pulse train)
        M_echomat=zeros(nEcho,echosamples); %matrix storing the acquired
            data (practically this is finally the k-space)

        %select phantom line
        tmp=squeeze(phant(li,:));
        M_0=squeeze(kron(tmp',ones(360,1))); %repeats every point of the
            line 360 times to make a quasi-continuous object
        M_mat(:,1)=M_0'; %line is imposed on equilibrium magnetization
```

```

% Excitation
M_mat(:,:)=pulse_function( M_mat(:,:),90,phasecyc(1) );
M_mat(:,:)=relax_function( M_0,M_mat(:,:),ES/2,T1,T2 );
M_mat(:,:)=dephase_function( M_mat(:,:),crushergrad,start,offres
    ); %dephasing from crusher

%Refocusing pulse train
for pulse=1:nEcho
    if any(pulse==nDiff) % apply diffusion weighting
        (stejskal-tanner)
        M_mat(:,:)=pulse_function( M_mat(:,:),alpha,phasecyc(2) );
        M_mat(:,:)=dephase_function(
            M_mat(:,:),diffgrad,start,offres ); %dephasing from
            diffusion gradient
        M_mat(:,:)=relax_function( M_0,M_mat(:,:),ES,T1,T2 );
        M_mat(:,:)=pulse_function( M_mat(:,:),alpha,phasecyc(2) );
        M_mat(:,:)=dephase_function(
            M_mat(:,:),diffgrad,start,offres );
        M_mat(:,:)=relax_function( M_0,M_mat(:,:),ES,T1,T2 );
    end
    % Read module
    M_mat(:,:)=pulse_function( M_mat(:,:),alpha,phasecyc(pulse+1)
        );
    M_mat(:,:)=dephase_function(
        M_mat(:,:),crushergrad,start,offres ); %rephasing from
        crusher
    M_mat(:,:)=dephase_function(
        M_mat(:,:),-readgrad,start,offres ); %read prephaser
    for k=1:echosamples % ADC block
        M_mat(:,:)=dephase_function(
            M_mat(:,:),2*readgrad/echosamples,start,offres );
        M_mat(:,:)=relax_function(
            M_0,M_mat(:,:),ES/echosamples,T1,T2 );
        M_echomat(pulse,k)=squeeze( sum(M_mat(:,2)) );
    end
    M_mat(:,:)=dephase_function(
        M_mat(:,:),-readgrad,start,offres ); %read rephaser
    M_mat(:,:)=dephase_function(
        M_mat(:,:),crushergrad,start,offres ); %dephasing from
        crusher
end
%make phase encoding (direct center-out order)
for s=1:nEcho
    linespace(li,rep,s,:)=squeeze(M_echomat(s,:))
        .*exp(2*pi*1j*(li-1)*(rep-1)/(nX));
end

```

```

    end %eof line loop
end %eof repetition loop
Echoes=linekspace; %return this matrix

%functions
function M_back=relax_function(M_0,M,t,T1,T2); %function for T1 and T2
relaxation
bloch=[exp(-t/T1) 0; 0 exp(-t/T2)];
z_mat=zeros(size(M));
z_mat(:,1)=M_0*(1-exp(-t/T1));
M_back=M*bloch+z_mat;

function M_back=dephase_function(M,deph_fraction,start,offres); %function
of dephasing per TR, assuming 360 isochromats per voxel
M_back=zeros(size(M));
dephmat=[ones(size(M,1),1) exp(-j*offres*pi/180)
*exp(-j*[1-start:size(M,1)-start]' *deph_fraction*2*pi/(360)
)];%size(M,1) ];
M_back=M.*dephmat;

function M_back=pulse_function(M,angle,phase); %function for RF-pulse
% see Liang/Lauterbur page 82
M_temp=[real(M(:,2)) imag(M(:,2)) M(:,1)];
phase=phase*pi/180;
angle=angle*pi/180;
puls_temp=[(cos(angle)*sin(phase)*sin(phase)+cos(phase)*cos(phase))
(sin(angle/2)*sin(angle/2)*sin(2*phase)) - (sin(angle)*sin(phase));
(sin(angle/2)*sin(angle/2)*sin(2*phase))
(sin(phase)*sin(phase)+cos(angle)*cos(phase)*cos(phase))
(sin(angle)*cos(phase));
(sin(angle)*sin(phase)) - (cos(phase)*sin(angle)) (cos(angle));];
M_back_temp=M_temp*puls_temp';
M_back=[M_back_temp(:,3) M_back_temp(:,1)+j*M_back_temp(:,2)];

```

Listing 11.4: Parameter setup for the in Vivo measurement with the SATM Sequence.

```

param_seq.TR          = 5000e-3; % s
param_seq.nX          = 128; %matrix size
param_seq.nEcho       = 30;
param_seq.nSlices     = 1;
param_seq.nSpokes     = param_seq.nEcho*30;
param_seq.nDiffusionEcho = [1 6 11 16];

param_seq.fov         = 250e-3; % m
param_seq.sliceThickness = 5e-3; % m
param_seq.sliceDist   = 1*param_seq.sliceThickness; % m

param_seq.dG          = 150e-6; % s
param_seq.dGdiff      = 1500e-6; % s
param_seq.tdiffDead   = 0e-3; % s
param_seq.tEx         = 3e-3; % s

param_seq.TE          = [9, 9.5, 10, 12]*1e-3; % s
param_seq.readoutTime = 4.8e-3; % s

%%
param_seq.invert_enc   = false;
param_seq.disable_radial = false;
param_seq.disable_diff = false;
param_seq.cartesian_enc = false;
param_seq.scanner     = 'prisma';
param_seq.diff        = 'z';
%%
param_seq.verbose     = 1;
param_seq.display     = 0;
param_seq.write       = 1;
%%
param_seq.filen = ['Experiments', filesep, '20210629_rachel', filesep,
                  'setup', filesep, 'InVivo30'];
[seq,param_sim] = main_seq(param_seq);

```

BIBLIOGRAPHY

- [Akk17] Akkus, Z., Galimzianova, A., Hoogi, A., et al. **Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions**. In: *Journal of Digital Imaging* 30.4 (Aug. 2017), pp. 449–459.
- [Ali18] Aliotta, E., Moulin, K., Zhang, Z., et al. **Simultaneous Measurement of T2 and Apparent Diffusion Coefficient (T2+ADC) in the Heart With Motion-Compensated Spin Echo Diffusion-Weighted Imaging**. In: *Magnetic resonance in medicine* 79.2 (Feb. 2018), pp. 654–662.
- [Ang10] Ang, K. K., Harris, J., Wheeler, R., et al. **Human Papillomavirus and Survival of Patients with Oropharyngeal Cancer**. In: *The New England Journal of Medicine* 363.1 (July 2010), pp. 24–35.
- [AHI09] Atcheson, B., Heidrich, W., and Ihrke, I. **An Evaluation of Optical Flow Algorithms for Background Oriented Schlieren Imaging**. In: *Experiments in Fluids* 46.3 (Mar. 2009), pp. 467–476.
- [Bak17] Bakas, S., Akbari, H., Sotiras, A., et al. **Segmentation Labels for the Pre-operative Scans of the TCGA-LGG Collection**. 2017.
- [BKZ04] Bernstein, M. A., King, K. F., and Zhou, Z. J. **Handbook of MRI Pulse Sequences**. Amsterdam ; Boston: Academic Press, 2004.
- [BB17] Bielak, L. and Bock, M. **Optimization of Diffusion Imaging for Multiple Target Regions Using Maximum Likelihood Estimation**. In: *Current Directions in Biomedical Engineering* 3.2 (Sept. 2017), pp. 203–206.
- [Bie18a] Bielak, L., Bunea, H., Wiedenmann, N., et al. **Image-Based Multi-Scale Distortion Correction: Application to Diffusion Imaging**. In: *Proc. Int. Soc. Mag. Reson. Med.* Vol. 26. Paris, France, June 2018, p. 1633.
- [Bie18b] Bielak, L., Bunea, H., Wiedenmann, N., et al. **The Weakest Link in the Chain: How MR Data Quality Influences Convolutional Neural Network Performance**. In: *Proc. Int. Soc. Mag. Reson. Med.* Vol. 26. Paris, France, June 2018, p. 2720.
- [BLB21] Bielak, L., Lottner, T., and Bock, M. **Radial Fast Spin Echo MRI with Compressed Sensing for Simultaneous ADC and T2 Mapping**. In: *Proc. Int. Soc. Mag. Reson. Med.* Vol. 29. Virtual Meeting, May 2021, p. 4166.

- [Bie20a] Bielał, L., Wiedenmann, N., Berlin, A., et al. **Evaluating Noise Robustness of CNN-based Head&Neck Tumor Segmentations on Multiparametric MRI Data**. In: *Proc. Int. Soc. Mag. Reson. Med.* Vol. 28. Virtual Meeting, Aug. 2020, p. 0269.
- [Bie20b] Bielał, L., Wiedenmann, N., Berlin, A., et al. **Convolutional Neural Networks for Head and Neck Tumor Segmentation on 7-Channel Multiparametric MRI: A Leave-One-out Analysis**. In: *Radiation Oncology* 15.1 (July 2020), p. 181.
- [Bie19a] Bielał, L., Wiedenmann, N., Lottner, T., et al. **Quantifying Information Content of Multiparametric MRI Data for Automatic Tumor Segmentation Using CNNs**. In: *Proc. Int. Soc. Mag. Reson. Med.* Vol. 27. Montréal, QC, Canada, Apr. 2019, p. 2339.
- [Bie19b] Bielał, L., Wiedenmann, N., Nicolay, N. H., et al. **Automatic Tumor Segmentation With a Convolutional Neural Network in Multiparametric MRI: Influence of Distortion Correction**. In: *Tomography: A Journal for Imaging Research* 5.3 (Sept. 2019), pp. 292–299.
- [BRN96] Blamire, A. M., Rothman, D. L., and Nixon, T. **Dynamic Shim Updating: A New Approach towards Optimized Whole Brain Shimming**. In: *Magnetic Resonance in Medicine* 36.1 (July 1996), pp. 159–165.
- [BUF09] Block, K., Uecker, M., and Frahm, J. **Model-Based Iterative Reconstruction for Radial Fast Spin-Echo MRI**. In: *IEEE Transactions on Medical Imaging* 28.11 (Nov. 2009), pp. 1759–1769.
- [BUF07] Block, K. T., Uecker, M., and Frahm, J. **Undersampled Radial MRI with Multiple Coils. Iterative Image Reconstruction Using a Total Variation Constraint**. In: *Magnetic Resonance in Medicine* 57.6 (June 2007), pp. 1086–1098.
- [Blo88] Blot, W. J., McLaughlin, J. K., Winn, D. M., et al. **Smoking and Drinking in Relation to Oral and Pharyngeal Cancer**. In: *Cancer Research* 48.11 (June 1988), pp. 3282–3287.
- [BE84] Bottomley, P. A. and Edelstein, W. A. Method of Eliminating Effects of Spurious Free Induction Decay NMR Signal Caused by Imperfect 180 Degrees Pulses. In: US Patent 4484138 (Nov. 1984).
- [Bou00] Bouguet, J.-y. Pyramidal Implementation of the Lucas Kanade Feature Tracker. In: *Intel Corporation, Microprocessor Research Labs* (2000).
- [Bro14] Brown, R. W., Cheng, Y.-C. N., Haacke, E. M., et al. **Magnetic Resonance Imaging: Physical Principles and Sequence Design**. 2nd edition. Hoboken, New Jersey: John Wiley & Sons, Inc, 2014.

- [Buro9] Burnside, E. S., Sickles, E. A., Bassett, L. W., et al. **The ACR BI-RADS® Experience: Learning From History**. In: *Journal of the American College of Radiology* 6.12 (Dec. 2009), pp. 851–860.
- [CP54] Carr, H. Y. and Purcell, E. M. **Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments**. In: *Physical Review* 94.3 (May 1954), pp. 630–638.
- [Chu18] Chuter, R. W., Pollitt, A., Whitehurst, P., et al. **Assessing MR-linac Radiotherapy Robustness for Anatomical Changes in Head and Neck Cancer**. In: *Physics in Medicine & Biology* 63.12 (June 2018), p. 125020.
- [CS18] Connolly, M. and Srinivasan, A. **Diffusion-Weighted Imaging in Head and Neck Cancer: Technique, Limitations, and Applications**. In: *Magnetic Resonance Imaging Clinics of North America* 26.1 (Feb. 2018), pp. 121–133.
- [Des12] Deshmane, A., Gulani, V., Griswold, M. A., et al. **Parallel MR Imaging**. In: *Journal of Magnetic Resonance Imaging* 36.1 (July 2012), pp. 55–72.
- [DIC21] DICOM. **DICOM PS3.3 2021b - Information Object Definitions**. NEMA. 2021. URL: <http://dicom.nema.org/medical/dicom/current/output/chtml/part03/PS3.3.html> (visited on 06/07/2021).
- [Dono6] Donoho, D. **Compressed Sensing**. In: *IEEE Transactions on Information Theory* 52.4 (Apr. 2006), pp. 1289–1306.
- [Dri15] Driessen, J. P., Kempen, P. M. W. van, Heijden, G. J. van der, et al. **Diffusion-Weighted Imaging in Head and Neck Squamous Cell Carcinomas: A Systematic Review**. In: *Head & Neck* 37.3 (Mar. 2015), pp. 440–448.
- [Duy98] Duyn, J. H., Yang, Y., Frank, J. A., et al. **Simple Correction Method Fork-Space Trajectory Deviations in MRI**. In: *Journal of Magnetic Resonance* 132.1 (May 1998), pp. 150–153.
- [Eino5] Einstein, A. **Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen**. In: *Annalen der Physik* 322.8 (1905), pp. 549–560.
- [Fei86] Feinberg, D. A., Hale, J. D., Watts, J. C., et al. **Halving MR Imaging Time by Conjugation: Demonstration at 3.5 kG**. In: *Radiology* 161.2 (Nov. 1986), pp. 527–531.
- [Fes] Fessler, J. A. **Michigan Image Reconstruction Toolbox**. URL: <https://github.com/JeffFessler/mirt>.

- [Fes07] Fessler, J. A. **On NUFFT-based Gridding for Non-Cartesian MRI**. In: *Journal of Magnetic Resonance* 188.2 (Oct. 2007), pp. 191–195.
- [Fic55] Fick, A. **Ueber Diffusion**. In: *Annalen der Physik und Chemie* 170.1 (1855), pp. 59–86.
- [Gat16] Gatidis, S., Graf, H., Weiß, J., et al. **Diffusion-Weighted Echo Planar MR Imaging of the Neck at 3 T Using Integrated Shimming: Comparison of MR Sequence Techniques for Reducing Artifacts Caused by Magnetic-Field Inhomogeneities**. In: *Magnetic Resonance Materials in Physics, Biology and Medicine* (Aug. 2016), pp. 1–7.
- [Gat14] Gatidis, S., Schmidt, H., Martirosian, P., et al. **Development of an MRI Phantom for Diffusion-Weighted Imaging with Independent Adjustment of Apparent Diffusion Coefficient Values and T2 Relaxation Times**. In: *Magnetic Resonance in Medicine* 72.2 (Aug. 2014), pp. 459–463.
- [GJC01] Gerig, G., Jomier, M., and Chakos, M. **Valmet: A New Validation Tool for Assessing and Improving 3D Object Segmentation**. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2001*. Ed. by W. J. Niessen and M. A. Viergever. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2001, pp. 516–523.
- [GB10] Glorot, X. and Bengio, Y. **Understanding the Difficulty of Training Deep Feedforward Neural Networks**. In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Mar. 2010, pp. 249–256.
- [GP06] Gourin, C. G. and Podolsky, R. H. **Racial Disparities in Patients with Head and Neck Squamous Cell Carcinoma**. In: *The Laryngoscope* 116.7 (July 2006), pp. 1093–1106.
- [Gra17] Gras, V., Farrher, E., Grinberg, F., et al. **Diffusion-Weighted DESS Protocol Optimization for Simultaneous Mapping of the Mean Diffusivity, Proton Density and Relaxation Times at 3 Tesla**. In: *Magnetic Resonance in Medicine* 78.1 (2017), pp. 130–141.
- [GK09] Greenberg, D. S. and Kerr, J. N. D. **Automated Correction of Fast Motion Artifacts for Two-Photon Imaging of Awake Animals**. In: *Journal of Neuroscience Methods* 176.1 (Jan. 2009), pp. 1–15.
- [Gri02] Griswold, M. A., Jakob, P. M., Heidemann, R. M., et al. **Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA)**. In: *Magnetic Resonance in Medicine* 47.6 (June 2002), pp. 1202–1210.

- [GN15] Grosu, A.-L. and Nicolay, N. H. **Evaluation of ^{18}F -MISO-PET Detectable Changes in Proliferation and Dispersion of Hypoxic Subvolume in Head and Neck Cancer in Patients Undergoing Radio(Chemo)Therapy.** German Clinical Trial Register. DRKS-ID: 00003830. Federal Institute for Drugs and Medical Devices, Cologne, Germany. Aug. 2015. URL: https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00003830.
- [GP95] Gudbjartsson, H. and Patz, S. **The Rician Distribution of Noisy MRI Data.** In: *Magnetic Resonance in Medicine* 34.6 (Dec. 1995), pp. 910–914.
- [Gue19] Guezenec, C., Bourhis, D., Orhac, F., et al. **Inter-Observer and Segmentation Method Variability of Textural Analysis in Pre-Therapeutic FDG PET/CT in Head and Neck Cancer.** In: *PLOS ONE* 14.3 (Mar. 2019). Ed. by A. D. Hutson, e0214299.
- [HLL91] Haacke, E., Lindskog, E., and Lin, W. **A Fast, Iterative, Partial-Fourier Technique Capable of Local Phase Recovery.** In: *Journal of Magnetic Resonance (1969)* 92.1 (Mar. 1991), pp. 126–145.
- [HZ05] Hager, W. and Zhang, H. **A New Conjugate Gradient Method with Guaranteed Descent and an Efficient Line Search.** In: *SIAM J. Optim.* (2005).
- [Hah50] Hahn, E. L. **Spin Echoes.** In: *Physical Review* 80.4 (Nov. 1950), pp. 580–594.
- [Han17] Hancu, I., Lee, S.-K., Hulsey, K., et al. **Distortion Correction in Diffusion-Weighted Imaging of the Breast: Performance Assessment of Prospective, Retrospective, and Combined (Prospective + Retrospective) Approaches.** In: *Magnetic Resonance in Medicine* 78.1 (2017), pp. 247–253.
- [HST10] Harari, P. M., Song, S., and Tomé, W. A. **Emphasizing Conformal Avoidance Versus Target Definition for IMRT Planning in Head-and-Neck Cancer.** In: *International Journal of Radiation Oncology*Biography*Physics* 77.3 (July 2010), pp. 950–958.
- [Hav17] Havaei, M., Davy, A., Warde-Farley, D., et al. **Brain Tumor Segmentation with Deep Neural Networks.** In: *Medical Image Analysis* 35 (Jan. 2017), pp. 18–31.
- [He15] He, K., Zhang, X., Ren, S., et al. **Deep Residual Learning for Image Recognition.** In: *arXiv:1512.03385 [cs]* (Dec. 2015). arXiv: [1512.03385 \[cs\]](https://arxiv.org/abs/1512.03385).

- [HNF86] Hennig, J., Nauerth, A., and Friedburg, H. **RARE Imaging: A Fast Imaging Method for Clinical MR**. In: *Magnetic Resonance in Medicine* 3.6 (Dec. 1986), pp. 823–833.
- [Hin12] Hinton, G. E. **A Practical Guide to Training Restricted Boltzmann Machines**. In: *Neural Networks: Tricks of the Trade: Second Edition*. Ed. by G. Montavon, G. B. Orr, and K.-R. Müller. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2012, pp. 599–619.
- [Hino2] Hinton, G. E. **Training Products of Experts by Minimizing Contrastive Divergence**. In: *Neural Computation* 14.8 (Aug. 2002), pp. 1771–1800.
- [HSW89] Hornik, K., Stinchcombe, M., and White, H. **Multilayer Feedforward Networks Are Universal Approximators**. In: *Neural Networks* 2.5 (Jan. 1989), pp. 359–366.
- [IS15] Ioffe, S. and Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In: *arXiv: 1502.03167 [cs]* (Mar. 2015). arXiv: [1502.03167 \[cs\]](https://arxiv.org/abs/1502.03167).
- [Jac91] Jackson, J., Meyer, C., Nishimura, D., et al. **Selection of a Convolution Function for Fourier Inversion Using Gridding (Computerised Tomography Application)**. In: *IEEE Transactions on Medical Imaging* 10.3 (Sept. 1991), pp. 473–478.
- [Jac19] Jacobsen, N., Deistung, A., Timmann, D., et al. **Analysis of Intensity Normalization for Optimal Segmentation Performance of a Fully Convolutional Neural Network**. In: *Zeitschrift für Medizinische Physik*. Special Issue: Deep Learning in Medical Physics 29.2 (May 2019), pp. 128–138.
- [JRS16] Jha, A. K., Rodríguez, J. J., and Stopeck, A. T. **A Maximum-Likelihood Method to Estimate a Single ADC Value of Lesions Using Diffusion MRI**. In: *Magnetic Resonance in Medicine* 76.6 (Dec. 2016), pp. 1919–1931.
- [Jon10] D. K. Jones, ed. **Diffusion MRI: Theory, Methods, and Application**. Oxford ; New York: Oxford University Press, 2010.
- [Kam17] Kamnitsas, K., Ledig, C., Newcombe, V. F. J., et al. **Efficient Multi-Scale 3D CNN with Fully Connected CRF for Accurate Brain Lesion Segmentation**. In: *Medical Image Analysis* 36 (Feb. 2017), pp. 61–78.
- [Kat] Kataria, T., Gupta, D., Goyal, S., et al. **Clinical Outcomes of Adaptive Radiotherapy in Head and Neck Cancers**. In: *The British Journal of Radiology* 89.1062 ().

- [KB17] Kingma, D. P. and Ba, J. Adam: A Method for Stochastic Optimization. In: *arXiv:1412.6980 [cs]* (Jan. 2017). arXiv: [1412.6980 \[cs\]](#).
- [KSÖ07] Koay, C. G., Sarlls, J. E., and Özarlan, E. **Three-Dimensional Analytical Magnetic Resonance Imaging Phantom in the Fourier Domain**. In: *Magnetic Resonance in Medicine* 58.2 (2007), pp. 430–436.
- [KK11] Krähenbühl, P. and Koltun, V. Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials. In: *Advances in Neural Information Processing Systems*. Ed. by J. Shawe-Taylor, R. Zemel, P. Bartlett, et al. Vol. 24. Curran Associates, Inc., 2011.
- [KK12] Krähenbühl, P. and Koltun, V. Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials. In: *arXiv:1210.5644 [cs]* (Oct. 2012). arXiv: [1210.5644 \[cs\]](#).
- [KSH12] Krizhevsky, A., Sutskever, I., and Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. In: *Advances in Neural Information Processing Systems*. Ed. by F. Pereira, C. J. C. Burges, L. Bottou, et al. Vol. 25. Curran Associates, Inc., 2012.
- [KSH17] Krizhevsky, A., Sutskever, I., and Hinton, G. E. **ImageNet Classification with Deep Convolutional Neural Networks**. In: *Communications of the ACM* 60.6 (May 2017), pp. 84–90.
- [KN08] Kuroki, Y. and Nasu, K. **Advances in Breast MRI: Diffusion-Weighted Imaging of the Breast**. In: *Breast Cancer* 15.3 (July 2008), pp. 212–217.
- [KEN91] Kwiat, D., Einav, S., and Navon, G. **A Decoupled Coil Detector Array for Fast Image Acquisition in Magnetic Resonance Imaging: Technical Report: Decoupled Coil Detector Array**. In: *Medical Physics* 18.2 (Mar. 1991), pp. 251–265.
- [Lau73] Lauterbur, P. C. **Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance**. In: *Nature* 242.5394 (Mar. 1973), pp. 190–191.
- [Lay17] Layton, K. J., Kroboth, S., Jia, F., et al. **Pulseq: A Rapid and Hardware-Independent Pulse Sequence Prototyping Framework**. In: *Magnetic Resonance in Medicine* 77.4 (2017), pp. 1544–1552.
- [Le 86] Le Bihan, D., Breton, E., Lallemand, D., et al. **MR Imaging of Intravoxel Incoherent Motions: Application to Diffusion and Perfusion in Neurologic Disorders**. In: *Radiology* 161.2 (Nov. 1986), pp. 401–407.
- [Lec98] Lecun, Y., Bottou, L., Bengio, Y., et al. **Gradient-Based Learning Applied to Document Recognition**. In: *Proceedings of the IEEE* 86.11 (Nov./1998), pp. 2278–2324.

- [LBH15] LeCun, Y., Bengio, Y., and Hinton, G. **Deep Learning**. In: *Nature* 521.7553 (May 2015), pp. 436–444.
- [LeC12] LeCun, Y. A., Bottou, L., Orr, G. B., et al. **Efficient BackProp**. In: *Neural Networks: Tricks of the Trade: Second Edition*. Ed. by G. Montavon, G. B. Orr, and K.-R. Müller. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2012, pp. 9–48.
- [Levo8] Levitt, M. H. **Spin Dynamics: Basics of Nuclear Magnetic Resonance**. 2nd ed. Chichester, England ; Hoboken, NJ: John Wiley & Sons, 2008.
- [Lin17] Lin, T.-Y., Goyal, P., Girshick, R., et al. **Focal Loss for Dense Object Detection**. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. Oct. 2017, pp. 2999–3007.
- [Lit17a] Litjens, G., Debats, O., Barentsz, J., et al. **SPIE-AAPM PROSTATEx Challenge Data**. 2017.
- [Lit17b] Litjens, G., Kooi, T., Bejnordi, B. E., et al. **A Survey on Deep Learning in Medical Image Analysis**. In: *Medical Image Analysis* 42 (Dec. 2017), pp. 60–88.
- [LK81] Lucas, B. D. and Kande, T. An Iterative Image Registration Technique with an Application to Stereo Vision. In: *Proceedings of the 1981 DARPA Image Understanding Workshop*. Apr. 1981, pp. 121–130.
- [MPV88] MacFall, J. R., Pelc, N. J., and Vavrek, R. M. **Correction of Spatially Dependent Phase Shifts for Partial Fourier Imaging**. In: *Magnetic Resonance Imaging* 6.2 (Mar. 1988), pp. 143–155.
- [MSK21] Magny, S. J., Shikhman, R., and Keppke, A. L. Breast Imaging Reporting and Data System. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2021.
- [Man77] Mansfield, P. **Multi-Planar Image Formation Using NMR Spin Echoes**. In: 10.3 (Feb. 1977), pp. L55–L58.
- [McG93] McGibney, G., Smith, M. R., Nichols, S. T., et al. **Quantitative Evaluation of Several Partial Fourier Reconstruction Algorithms Used in Mri**. In: *Magnetic Resonance in Medicine* 30.1 (July 1993), pp. 51–59.
- [MDL17] Men, K., Dai, J., and Li, Y. **Automatic Segmentation of the Clinical Target Volume and Organs at Risk in the Planning CT for Rectal Cancer Using Deep Dilated Convolutional Neural Networks**. In: *Medical Physics* 44.12 (Dec. 2017), pp. 6377–6389.
- [Mey93] Meyer, Y. **Wavelets and Operators**. Trans. by D. H. Salinger. First. Cambridge University Press, Apr. 1993.

- [MLB15] Mispelter, J., Lupu, M., and Briguet, A. **NMR Probeheads for Biophysical and Biomedical Experiments: Theoretical Principles & Practical Guidelines**. Second edition. London: Imperial College Press, 2015.
- [Moe19] Moe, Y. M., Groendahl, A. R., Mulstad, M., et al. Deep Learning for Automatic Tumour Segmentation in PET/CT Images of Patients with Head and Neck Cancers. In: *arXiv:1908.00841 [eess]* (Aug. 2019). arXiv: [1908.00841 \[eess\]](https://arxiv.org/abs/1908.00841).
- [MS97] Morrell, G. and Spielman, D. **Dynamic Shimming for Multi-Slice Magnetic Resonance Imaging**. In: *Magnetic Resonance in Medicine* 38.3 (Sept. 1997), pp. 477–483.
- [MCL13] Morris, E. A., Comstock, C. E., and Lee, C. H. **ACR BI-RADS® Magnetic Resonance Imaging**. In: *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*. Reston, VA: American College of Radiology, 2013.
- [Mou14] Moussavi, A., Untenberger, M., Uecker, M., et al. **Correction of Gradient-Induced Phase Errors in Radial MRI**. In: *Magnetic Resonance in Medicine* 71.1 (2014), pp. 308–312.
- [Nag11] Nagi, J., Ducatelle, F., Di Caro, G. A., et al. **Max-Pooling Convolutional Neural Networks for Vision-Based Hand Gesture Recognition**. In: *2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*. Nov. 2011, pp. 342–347.
- [New21] Newitt, D. C., Partridge, S. C., Zhang, Z., et al. **ACRIN 6698/I-SPY2 Breast DWI**. 2021.
- [Nil10] Nilsson, N. J. **The Quest for Artificial Intelligence: A History of Ideas and Achievements**. Cambridge ; New York: Cambridge University Press, 2010.
- [Oga90] Ogawa, S., Lee, T. M., Kay, A. R., et al. **Brain Magnetic Resonance Imaging with Contrast Dependent on Blood Oxygenation**. In: *Proceedings of the National Academy of Sciences* 87.24 (Jan. 1990), pp. 9868–9872.
- [Oud16] Oudeman, J., Coolen, B. F., Mazzoli, V., et al. **Diffusion-Prepared Neurography of the Brachial Plexus with a Large Field-of-View at 3T**. In: *Journal of Magnetic Resonance Imaging* 43.3 (Mar. 2016), pp. 644–654.

- [Pau91] Pauly, J., Le Roux, P., Nishimura, D., et al. **Parameter Relations for the Shinnar-Le Roux Selective Excitation Pulse Design Algorithm (NMR Imaging)**. In: *IEEE Transactions on Medical Imaging* 10.1 (Mar. 1991), pp. 53–65.
- [Per15] Pereira, S., Pinto, A., Alves, V., et al. **Deep Convolutional Neural Networks for the Segmentation of Gliomas in Multi-sequence MRI**. In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Lecture Notes in Computer Science. Springer, Cham, Oct. 2015, pp. 131–143.
- [PW17] Perez, L. and Wang, J. The Effectiveness of Data Augmentation in Image Classification Using Deep Learning. In: *arXiv:1712.04621 [cs]* (Dec. 2017). arXiv: 1712.04621 [cs].
- [PDM03] Peters, D. C., Derbyshire, J. A., and McVeigh, E. R. **Centering the Projection Reconstruction Trajectory: Reducing Gradient Delay Errors**. In: *Magnetic Resonance in Medicine* 50.1 (2003), pp. 1–6.
- [Pfa18] Pfaffenrot, V., Brunheim, S., Rietsch, S. H. G., et al. **An 8/15-Channel Tx/Rx Head Neck RF Coil Combination with Region-Specific B₁+ Shimming for Whole-Brain MRI Focused on the Cerebellum at 7T**. In: *Magnetic Resonance in Medicine* 80.3 (2018), pp. 1252–1265.
- [PM99] Pipe, J. G. and Menon, P. **Sampling Density Compensation in MRI: Rationale and an Iterative Numerical Solution**. In: *Magnetic Resonance in Medicine* 41.1 (1999), pp. 179–186.
- [PG17] Pnevmatikakis, E. A. and Giovannucci, A. **NoRMCorre: An Online Algorithm for Piecewise Rigid Motion Correction of Calcium Imaging Data**. In: *Journal of Neuroscience Methods* 291 (Nov. 2017), pp. 83–94.
- [Pru99] Pruessmann, K. P., Weiger, M., Scheidegger, M. B., et al. **SENSE: Sensitivity Encoding for Fast MRI**. In: *Magnetic Resonance in Medicine* 42.5 (Nov. 1999), pp. 952–962.
- [Pus88] Pusey, E., Yoon, C., Anselmo, M. L., et al. **Aliasing Artifacts in Mr Imaging**. In: *Computerized Medical Imaging and Graphics* 12.4 (July 1988), pp. 219–224.
- [Qiu21] Qiu, J., Liu, J., Bi, Z., et al. **Integrated Slice-Specific Dynamic Shimming Diffusion Weighted Imaging (DWI) for Rectal Cancer Detection and Characterization**. In: *Cancer Imaging* 21.1 (Apr. 2021), p. 32.

- [Rac21] Rachel, T. Parameteroptimierung Für Die Simultane Messung Der T₂-Relaxation Und Des Diffusionskoeffizienten in Radialen, Unterabgetasteten Spin-Echo-Sequenzen. Bachelor Thesis. Fakultät für Mathematik und Physik: Albert-Ludwigs-Universität Freiburg, June 2021.
- [Rau11] Rauschenberg, J., Krafft, A. J., Maier, F., et al. **Outer Volume Suppression in Steady State Sequences (OVSuSS) for Percutaneous Interventions**. In: *Magnetic Resonance in Medicine* 66.1 (2011), pp. 123–134.
- [RSH08] M. Reiser, W. Semmler, and H. Hricak, eds. **Magnetic Resonance Tomography**. Berlin: Springer, 2008.
- [RM51] Robbins, H. and Monro, S. **A Stochastic Approximation Method**. In: *The Annals of Mathematical Statistics* 22.3 (Sept. 1951), pp. 400–407.
- [Roeg90] Roemer, P. B., Edelstein, W. A., Hayes, C. E., et al. **The NMR Phased Array**. In: *Magnetic Resonance in Medicine* 16.2 (Nov. 1990), pp. 192–225.
- [RFB15] Ronneberger, O., Fischer, P., and Brox, T. **U-Net: Convolutional Networks for Biomedical Image Segmentation**. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by N. Navab, J. Hornegger, W. M. Wells, et al. Lecture Notes in Computer Science. Springer International Publishing, 2015, pp. 234–241.
- [RHU19] Rosenzweig, S., Holme, H. C. M., and Uecker, M. **Simple Auto-calibrated Gradient Delay Estimation from Few Spokes Using Radial Intersections (RING)**. In: *Magnetic Resonance in Medicine* 81.3 (Mar. 2019), pp. 1898–1906.
- [ROF92] Rudin, L. I., Osher, S., and Fatemi, E. **Nonlinear Total Variation Based Noise Removal Algorithms**. In: *Physica D: Nonlinear Phenomena* 60.1-4 (Nov. 1992), pp. 259–268.
- [RM86] Rumelhart, D. E. and McClelland, J. L. **Information Processing in Dynamical Systems: Foundations of Harmony Theory**. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*. MIT Press, 1986, pp. 194–281.
- [SLN11] Saritas, E. U., Lee, J. H., and Nishimura, D. G. **SNR Dependence of Optimal Parameters for Apparent Diffusion Coefficient Measurements**. In: *IEEE transactions on medical imaging* 30.2 (Feb. 2011), pp. 424–437.

- [SH98] Scheffler, K. and Hennig, J. **Reduced Circular Field-of-View Imaging**. In: *Magnetic Resonance in Medicine* 40.3 (Sept. 1998), pp. 474–480.
- [SG91] Schneider, E. and Glover, G. **Rapid in Vivo Proton Shimming**. In: *Magnetic Resonance in Medicine* 18.2 (1991), pp. 335–347.
- [Sel17] Selvaraju, R. R., Cogswell, M., Das, A., et al. **Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization**. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, Oct. 2017, pp. 618–626.
- [Sha20] Sharif, M. I., Li, J. P., Khan, M. A., et al. **Active Deep Neural Network Features Selection for Segmentation and Recognition of Brain Tumors Using MRI Images**. In: *Pattern Recognition Letters* 129 (Jan. 2020), pp. 181–189.
- [SL74] Shepp, L. A. and Logan, B. F. **The Fourier Reconstruction of a Head Section**. In: *IEEE Transactions on Nuclear Science* 21.3 (June 1974), pp. 21–43.
- [SZ15] Simonyan, K. and Zisserman, A. **Very Deep Convolutional Networks for Large-Scale Image Recognition**. In: *arXiv:1409.1556 [cs]* (Apr. 2015). arXiv: [1409.1556 \[cs\]](https://arxiv.org/abs/1409.1556).
- [SBE17] Skorpil, M., Brynolfsson, P., and Engström, M. **Motion Corrected DWI with Integrated T2-mapping for Simultaneous Estimation of ADC, T2-relaxation and Perfusion in Prostate Cancer**. In: *Magnetic Resonance Imaging* 39 (June 2017), pp. 162–167.
- [SLA12] Snoek, J., Larochelle, H., and Adams, R. P. **Practical Bayesian Optimization of Machine Learning Algorithms**. In: *arXiv:1206.2944 [cs, stat]* (June 2012). arXiv: [1206.2944 \[cs, stat\]](https://arxiv.org/abs/1206.2944).
- [Sri14] Srivastava, N., Hinton, G., Krizhevsky, A., et al. **Dropout: A Simple Way to Prevent Neural Networks from Overfitting**. In: *Journal of Machine Learning Research* 15 (2014), pp. 1929–1958.
- [Sta12] Staroswiecki, E., Granlund, K. L., Alley, M. T., et al. **Simultaneous Estimation of T2 and ADC in Human Articular Cartilage In Vivo with a Modified 3D DESS Sequence at 3 T**. In: *Magnetic Resonance in Medicine* 67.4 (Apr. 2012), pp. 1086–1096.
- [STM91] Stehling, M. K., Turner, R., and Mansfield, P. **Echo-Planar Imaging: Magnetic Resonance Imaging in a Fraction of a Second**. In: *Science* 254.5028 (Oct. 1991), pp. 43–50.
- [ST16] Steiger, P. and Thoeny, H. C. **Prostate MRI Based on PI-RADS Version 2: How We Review and Report**. In: *Cancer Imaging* 16 (2016), p. 9.

- [ST65] Stejskal, E. O. and Tanner, J. E. **Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient**. In: *The Journal of Chemical Physics* 42.1 (Jan. 1965), pp. 288–292.
- [Sud17] Sudre, C. H., Li, W., Vercauteren, T., et al. **Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations**. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Ed. by M. J. Cardoso, T. Arbel, G. Carneiro, et al. Vol. 10553. Cham: Springer International Publishing, 2017, pp. 240–248.
- [Sur17] Surucu, M., Shah, K. K., Roeske, J. C., et al. **Adaptive Radiotherapy for Head and Neck Cancer**. In: *Technology in Cancer Research & Treatment* 16.2 (Apr. 2017), pp. 218–223.
- [Suto5] Sutherland, W. **A Dynamical Theory of Diffusion for Non-Electrolytes and the Molecular Mass of Albumin**. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 9.54 (June 1905), pp. 781–785.
- [Suto4] Sutherland, W. **The Measurement of Large Molecular Masses**. In: *Report of the 10th Meeting of the Australasian Association for the Advancement of Science*. Vol. 10. Dunedin, New Zealand, 1904, pp. 117–121.
- [Sza21] Szandała, T. **Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks**. In: *Bio-Inspired Neurocomputing*. Ed. by A. K. Bhoi, P. K. Mallick, C.-M. Liu, et al. Studies in Computational Intelligence. Singapore: Springer, 2021, pp. 203–224.
- [TH15] Taha, A. A. and Hanbury, A. **Metrics for Evaluating 3D Medical Image Segmentation: Analysis, Selection, and Tool**. In: *BMC Medical Imaging* 15 (Aug. 2015), p. 29.
- [TK06] Theodoridis, S. and Koutroumbas, K. **Pattern Recognition**. 3rd ed. San Diego, CA: Academic Press, 2006.
- [Tof97] Tofts, P. S. **Modeling Tracer Kinetics in Dynamic Gd-DTPA MR Imaging**. In: *Journal of Magnetic Resonance Imaging* 7.1 (Jan. 1997), pp. 91–101.
- [Tof99] Tofts, P. S., Brix, G., Buckley, D. L., et al. **Estimating Kinetic Parameters from Dynamic Contrast-Enhanced T₁-Weighted MRI of a Diffusible Tracer: Standardized Quantities and Symbols**. In: *Journal of Magnetic Resonance Imaging* 10.3 (1999), pp. 223–232.
- [Tor56] Torrey, H. C. **Bloch Equations with Diffusion Terms**. In: *Physical Review* 104.3 (Nov. 1956), pp. 563–565.

- [Uec21] Uecker, M., Holme, C., Blumenthal, M., et al. **Mrirecon/Bart: Version 0.7.00**. Zenodo. Mar. 2021.
- [Uec14] Uecker, M., Lai, P., Murphy, M. J., et al. **ESPIRiT—an Eigenvalue Approach to Autocalibrating Parallel MRI: Where SENSE Meets GRAPPA**. In: *Magnetic Resonance in Medicine* 71.3 (Mar. 2014), pp. 990–1001.
- [van19] van der Veen, J., Willems, S., Deschuymer, S., et al. **Benefits of Deep Learning for Delineation of Organs at Risk in Head and Neck Cancer**. In: *Radiotherapy and Oncology* 138 (Sept. 2019), pp. 68–74.
- [vGN19] van der Veen, J., Gulyban, A., and Nuyts, S. **Interobserver Variability in Delineation of Target Volumes in Head and Neck Cancer**. In: *Radiotherapy and Oncology* 137 (Aug. 2019), pp. 9–15.
- [Vin14] Vinegoni, C., Lee, S., Feruglio, P. F., et al. **Advanced Motion Compensation Methods for Intravital Optical Microscopy**. In: *IEEE Journal of Selected Topics in Quantum Electronics* 20.2 (Mar. 2014), pp. 83–91.
- [VI97] Viola, P. and Iii, W. M. W. **Alignment by Maximization of Mutual Information**. In: *International Journal of Computer Vision* 24.2 (Sept. 1997), pp. 137–154.
- [Wah22] Wahid, K. A., Ahmed, S., He, R., et al. **Evaluation of Deep Learning-Based Multiparametric MRI Oropharyngeal Primary Tumor Auto-Segmentation and Investigation of Input Channel Effects: Results from a Prospective Imaging Registry**. In: *Clinical and Translational Radiation Oncology* 32 (Jan. 2022), pp. 6–14.
- [Wah21] Wahid, K. A., Ahmed, S., He, R., et al. **Development of a High-Performance Multiparametric MRI Oropharyngeal Primary Tumor Auto-Segmentation Deep Learning Model and Investigation of Input Channel Effects: Results from a Prospective Imaging Registry**. Preprint. *Radiology and Imaging*, July 2021.
- [Wal09] Walker-Samuel, S., Orton, M., McPhail, L. D., et al. **Robust Estimation of the Apparent Diffusion Coefficient (ADC) in Heterogeneous Solid Tumors**. In: *Magnetic Resonance in Medicine* 62.2 (Aug. 2009), pp. 420–429.
- [Wal17] Walter, S. S., Liu, W., Stemmer, A., et al. **Combination of Integrated Dynamic Shimming and Readout-Segmented Echo Planar Imaging for Diffusion Weighted MRI of the Head and Neck Region at 3Tesla**. In: *Magnetic Resonance Imaging* 42 (Oct. 2017), pp. 32–36.
- [Wan19] Wang, P. **On Defining Artificial Intelligence**. In: *Journal of Artificial General Intelligence* 10.2 (Jan. 2019), pp. 1–37.

- [WW20] Wang, Y. and Wang, M. **Selecting Proper Combination of mpMRI Sequences for Prostate Cancer Classification Using Multi-Input Convolutional Neuronal Network**. In: *Physica Medica* 80 (Dec. 2020), pp. 92–100.
- [Web16] A. Webb, ed. **Magnetic Resonance Technology: Hardware and System Component Design**. New Developments in NMR 7. Cambridge, UK: Royal Society of Chemistry, 2016.
- [Wei16] Weinreb, J. C., Barentsz, J. O., Choyke, P. L., et al. **PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2**. In: *European Urology* 69.1 (Jan. 2016), pp. 16–40.
- [Wie19] Wiedenmann, N., Bunea, H., Rischke, H., et al. **EP-2030 Multiparametric MRI and FMISO PET in HNSCC and Its Relation with Outcome**. In: *Radiotherapy and Oncology*. ESTRO 38, 26-30 April 2019, Milan, Italy 133 (Apr. 2019), S1114–S1115.
- [Wie18] Wiedenmann, N., Bunea, H., Rischke, H. C., et al. **Effect of Radiochemotherapy on T2* MRI in HNSCC and Its Relation to FMISO PET Derived Hypoxia and FDG PET**. In: *Radiat Oncol* 13 (Aug. 2018).
- [Wie20] Wiedenmann, N., Grosu, A.-L., Büchert, M., et al. **The Utility of Multiparametric MRI to Characterize Hypoxic Tumor Subvolumes in Comparison to FMISO PET/CT. Consequences for Diagnosis and Chemoradiation Treatment Planning in Head and Neck Cancer**. en. In: *Radiotherapy and Oncology* 150 (Sept. 2020), pp. 128–135.
- [Wino07] Winkelmann, S., Schaeffter, T., Koehler, T., et al. **An Optimal Radial Profile Order Based on the Golden Ratio for Time-Resolved MRI**. In: *IEEE transactions on medical imaging* 26.1 (Jan. 2007), pp. 68–76.
- [WB90] Wu, E. X. and Buxton, R. B. **Effect of Diffusion on the Steady-State Magnetization with Pulsed Field Gradients**. In: *Journal of Magnetic Resonance* (1969) 90.2 (Nov. 1990), pp. 243–253.
- [Wys13] Wyss, A., Hashibe, M., Chuang, S.-C., et al. **Cigarette, Cigar, and Pipe Smoking and the Risk of Head and Neck Cancers: Pooled Analysis in the International Head and Neck Cancer Epidemiology Consortium**. In: *American Journal of Epidemiology* 178.5 (Sept. 2013), pp. 679–690.
- [Yabo08] Yabuuchi, H., Matsuo, Y., Okafuji, T., et al. **Enhanced Mass on Contrast-Enhanced Breast MR Imaging: Lesion Characterization Using Combination of Dynamic Contrast-Enhanced and Diffusion-Weighted MR Images**. In: *Journal of Magnetic Resonance Imaging* 28.5 (Nov. 2008), pp. 1157–1165.

- [Yang98] Yan, D., Ziaja, E., Jaffray, D., et al. **The Use of Adaptive Radiation Therapy to Reduce Setup Error: A Prospective Clinical Study.** In: *International Journal of Radiation Oncology*Biology*Physics* 41.3 (June 1998), pp. 715-720.

*To struggle and to understand.
Never the last without the first.
That is the law.*

— George Mallory

DANKSAGUNG

Zuallererst möchte ich meiner großartigen Frau Sinja danken, dass sie mich kompromisslos unterstützt — und wann immer nötig auch gestützt — hat, in den langen Jahren die für diese Arbeit nötig waren.

Mein besonderer Dank gilt meinem Doktorvater Michael Bock, der eine herausragende Leistung als Betreuer gezeigt hat: Du hast genau das richtige Augenmaß für die Unterstützung, die hier und da notwendig ist, und der freien Forschung, die du uns Doktoranden lässt. Ich schätze mich sehr glücklich, dass mir deine Tür immer offen stand und dein Rat niemals hat lange auf sich warten lassen!

Ich freue mich sehr hier allen meinen Arbeitskollegen aus der Experimentellen Radiologie für die tolle Atmosphäre sowie die vielen kleinen und großen Hilfestellungen bei der Arbeit danken zu können. Insbesondere meinen langen Wegbegleitern Johannes Fischer und Thomas Lottner möchte ich danken, auch dafür, dass ihr während dieser Zeit zu engen Freunden wurdet. Ein großes Dankeschön geht auch an Deepa Gunashekar, denn ohne dich gäbe es kein Team in der ExpRad-KI-Gruppe! Vielen Dank für die Unterstützung und Beratung was die englische Sprache betrifft an meine *Mom* Karin sowie Samantha Hickey.

Die vielen Patientendaten in dieser Arbeit wären ohne Michaela Schäper niemals zustande gekommen, und ohne deine Mithilfe hätte ich die Feinheiten der Messungen so schnell nicht verstehen können — dafür einen herzlichen Dank!

Dieser Dank gilt auch den Kollegen aus der Strahlentherapie, die mich unterstützt und begleitet haben indem sie mir die klinische Sicht der Dinge näher brachten und die medizinischen Daten beigesteuert haben. Hier seien besonders Constantinos Zamboglou aus meiner Anfangszeit, sowie Nils Henrik Nicolay zu nennen, die einen großen Einsatz für mein Projekt gezeigt haben.

Natürlich geht mein Dank an Jürgen Hennig, der die Medizin Physik in meiner Zeit als Doktorand geführt hat und die Struktur schuf, ohne die meine Arbeit nicht möglich gewesen wäre.

Schlussendlich möchte ich meiner Familie danken, insbesondere meinen Eltern Elke und Eike. Dafür, dass ihr mir immer den Weg gewiesen habt, wenn ich nicht weiter wusste, ihr mir den Ehrgeiz mitgegeben habt, den Weg zu Ende zu gehen, so schwierig er manchmal auch zu sein schien, und ihr die Hand ausgestreckt habt, wenn die Hindernisse zu groß für mich schienen.

Dankeschön!