KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science

Unsupervised and semi-supervised Non-negative Matrix Factorization methods for brain tumor segmentation using multi-parametric MRI data

Nicolas Sauwen

Supervisors:
Prof. dr. ir. S. Van Huffel
Prof. dr. ir. F. Maes
Prof. dr. U. Himmelreich
Dr. D.M. Sima

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD)

December 2016

Unsupervised and semi-supervised Non-negative Matrix Factorization methods for brain tumor segmentation using multi-parametric MRI data

Nicolas SAUWEN

Examination committee:
Prof. dr. ir. P. Wollants, chair
Prof. dr. ir. S. Van Huffel, supervisor
Prof. dr. ir. F. Maes, supervisor
Prof. dr. U. Himmelreich, co-supervisor
Dr. D.M. Sima, co-supervisor
Prof. dr. ir. J. Suykens
Prof. dr. ir. K. Meerbergen
Prof. dr. ir. F. Glineur(UCL Louvain-La-Neuve)
Prof. dr. E. Achten(UZ Gent)
Dr. J. Slotboom(Inselspital Bern)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD)

December 2016

© 2016 KU Leuven – Faculty of Engineering Science
Uitgegeven in eigen beheer, Nicolas Sauwen, Kasteelpark Arenberg 10 box 2446, B-3001 Leuven (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Preface

In the first place I would like to thank my promotor, prof. Sabine Van Huffel, for giving me the opportunity to work on this PhD. Sabine, you have always pushed me to raise the bar a bit higher than I intended to, I won't deny that has been frustrating at times. But it never took me a long time to realize that it just improved the quality of my work. You should know that I very much enjoyed working on this PhD in the past 4 years, I know that matters to you. You never lose sight of the person that is behind the research, I admire that, especially in someone who has a many responsibilities as you do.

I can hardly thank enough my supervisor, dr. Diana Sima, for sticking with me on this winding road. Diana, switching to Dutch is no issue for you, so I want to tell you "dat ik met mijn gat in de boter ben gevallen". Thank you for being so patient, for always being supportive and for thinking along with me. I never felt like I was on my own in this work. Your advice has been invaluable for shaping this PhD.

I am also grateful for the help of my 2 other promotors, prof. Uwe Himmelreich and prof. Frederik Maes. Thank you both for your good advice during our regular meetings. I have appreciated your efforts for critically reviewing my work, always in a constructive way.

BioMed would not have been BioMed without all of my colleagues. It has been really great to be part of such a nice group, and to get to know so many kind and friendly people. Escaping heavy scientific thoughts was never really a problem during lunch (I will spare the reader from the disturbing topics that were often raised). In particular, I want to thank my brothers in arms, Adrian and Bharath, it will be weird not to have you guys around anymore every day. Bharath, thanks for always being so supportive and for the discussions about our work. Adrian, I know you claim to belong to the dark side, but you can't help shedding some light into most days. I will carry all of you with me along the journey and I can only hope to come across such nice co-workers once again.

I have had the pleasure to work together with some kind and helpful people during my PhD. I would like to thank dr. Sofie Van Cauter and dr. Marjan Acou for helping me with processing the MRI data, and for being patient with me while I sometimes tried to get into the radiologist's mind... Marjan, you really made our FWO project into an actual joint project. I have enjoyed our exchanges of ideas over the phone and I learnt a lot from it. I also want to thank dr. Jelle Veraart for helping our team with processing the diffusion-weighted MRI data, and for being so kind to share his code and knowledge so that we could become self-reliant.

I want to thank my friends for their support, especially during the tough summer months of thesis writing. We will never survive unless we get a little bit crazy... Faiza and Nick, thank you for the regular food deliveries during the writing months, such a nice gesture makes all the difference. Tom, Syb, Ingeborg, Wouter and Jurgen, thank you guys for your ongoing mental support and for pulling me away from the computer screen when I needed it.

Finally, a huge thank you to the best family in the world: moek, pap, Mil, Jan, Bo, thank you for always being there for me. When the going gets tough, you guys are always with me, even if you don't know it.

Abstract

Gliomas represent about 80% of all malignant primary brain tumors. Despite recent advancements in glioma research, patient outcome remains poor. The 5 year survival rate of the most common and most malignant subtype, i.e. glioblastoma, is about 5%. Magnetic resonance imaging (MRI) has become the imaging modality of choice in the management of brain tumor patients. Conventional MRI (cMRI) provides excellent soft tissue contrast without exposing the patient to potentially harmful ionizing radiation. Over the past decade, advanced MRI modalities, such as perfusion-weighted imaging (PWI), diffusion-weighted imaging (DWI) and magnetic resonance spectroscopic imaging (MRSI) have gained interest in the clinical field, and their added value regarding brain tumor diagnosis, treatment planning and follow-up has been recognized.

Tumor segmentation involves the imaging-based delineation of a tumor and its subcompartments. In gliomas, segmentation plays an important role in treatment planning as well as during follow-up. Manual segmentation by a clinical expert is currently the gold standard, but it is a tedious and time-consuming task. Gliomas are known to be heterogeneous: several stages of the disease can occur throughout the same lesion and diffuse boundaries exist between active tumor, necrosis, edema and the surrounding healthy brain. Clinical practice would benefit from accurate and automated volumetric delineation of the tumor and its subcompartments.

This thesis aims at developing methods for automated segmentation and characterization of brain tumors. The proposed methods are based on an unsupervised learning technique called Non-negative matrix factorization (NMF). NMF provides an additive parts-based representation of the input data, revealing the basic components which are present. Applied to the multi-parametric MR imaging data of a brain tumor patient, NMF is able to extract tissue-specific signatures as well as the relative proportions of the different tissue types in each voxel. Being an unsupervised method, NMF cannot benefit from an extensive training dataset to learn decision boundaries between tissue classes, but it is

directly applicable to any multi-parametric MRI (MP-MRI) dataset of any individual patient.

Two MP-MRI datasets were available for conducting the studies throughout this thesis. They were acquired at the university hospitals of Leuven (UZ Leuven) and Gent (UZ Gent), both including cMRI, PWI, DWI and MRSI, but with a different scanning protocol. Whereas only cMRI is usually considered for brain tumor segmentation, in this thesis all the available MP-MRI features are combined, and the added value of the individual MRI modalities will be explored.

A hierarchical variant of NMF called hierarchical NMF (hNMF) is presented, which essentially consists of 2 levels of NMF. hNMF is able to provide valid tissue differentiation, and segmentation performance in terms of the Dice score is competitive. When considering reduced MP-MRI datasets by omitting one MRI modality at a time, statistically significantly higher Dice scores were found for the high-grade glioma patients when using the full MP-MRI dataset. Computational efficiency of the hNMF algorithm is improved by considering advanced initialization methods. In particular, the successive projection algorithm (SPA) is proposed. Whereas SPA is commonly considered as a direct NMF source extraction tool, it was found in our work that better segmentation performance is achieved by using SPA as an initialization method. hNMF is also compared to other common unsupervised classifiers, including several single-level NMF and clustering methods. hNMF was found to outperform the other methods on the MP-MRI datasets of both hospitals. Thanks to its hierarchical structure, hNMF is better able to differentiate closely related tissue signatures, such as non-enhancing tumor and edema.

To be competitive with state-of-the-art, unsupervised segmentation algorithms require the incorporation of prior knowledge. A semi-supervised NMF-based segmentation framework is presented, incorporating user-defined voxel selection to initialize the NMF pathological tissue sources. L1-regularization is included in the NMF objective function to promote spatial consistency and sparseness of the tissue abundance maps. A morphological post-processing procedure exploits the location of the selected voxels along with tissue adjacency constraints to further enhance the segmentation results. Careful voxel selection in the tumor compartments was found necessary to obtain robust segmentation results. The semi-automated NMF-based segmentation framework was applied to the publicly available Leaderboard dataset of the BRATS 2013 challenge, allowing direct comparison with state-of-the-art. Competitive segmentation results are obtained, especially for the active tumor region for which our method ranks first among all participants. This thesis demonstrates that unsupervised segmentation algorithms can compete with state-of-the-art supervised segmentation methods by incorporating adequate prior knowledge into the algorithm.

Samenvatting

Gliomen vertegenwoordigen ongeveer 80% van alle kwaadaardige hersentumoren. Ondanks de recente vooruitgang in het onderzoek naar gliomen, blijven de vooruitzichten voor de patiënt slecht. Voor het meest voorkomende en meest kwaardaardige subtype, glioblastoma, bedraagt de overlevingskans na 5 jaar slechts 5%. Beeldvorming door magnetische resonantie (MRI) is de beeldmodaliteit bij voorkeur voor de opvolging van patiënten met een hersentumor. Conventionele MRI (cMRI) levert een uistekend contrast tussen de zachte weefseltypes, zonder de patiënt bloot te stellen aan schadelijke ioniserende straling. Gedurende het voorbije decennium hebben geavanceerde MRI modaliteiten zoals perfusie-gewogen beeldvorming (PWI), diffusie-gewogen beeldvorming (DWI), en magnetische resonantie spectroscopie beeldvorming (MRSI) aandacht verworven in de klinische praktijk. Hun toegevoegde waarde inzake de diagnose en de opvolging van hersentumoren is ondertussen erkend.

Tumor segmentatie betreft de aflijning van een tumor en zijn deelgebieden op basis van beelden. Segmentatie speelt een belangrijke rol bij de behandeling en de opvolging van gliomen. Manuele segmentatie wordt beschouwd als de gouden standaard, maar het is een moeilijke en tijdrovende taak. Gliomen zijn heterogeen: verschillende stadia van de ziekte kunnen voorkomen binnen dezelfde laesie en de grenzen tussen actieve tumor, necrose, oedeem en het gezond hersenweefsel zijn diffuus. De klinische praktijk zou baat hebben bij een nauwkeurige en geautomatiseerde aflijning van de tumor en zijn deelgebieden.

Deze thesis beoogt het ontwikkelen van methodes voor de automatische segmentatie en karakterisatie van hersentumoren. De methodes zijn gebaseerd op een ongesuperviseerde techniek genaamd Niet-negatieve matrix factorizatie (NMF). NMF levert een additieve onderdeel-gebaseerde voorstelling van de input data, en onthult daarbij de aanwezige basiscomponenten. Toegepast op de beelddata van een hersentumor patiënt, is NMF in staat om weefsel-specifieke signaturen af te leiden, alsook de relatieve proportie van de verschillende weefseltypes in elke voxel. Vermits NMF een ongesuperviseerde methode is,

kan ze geen beroep doen op uitgebreide training datasets om de weefseltypes correct te leren classificeren. NMF is echter wel rechtstreeks toepasbaar op elke multi-parametrische MRI (MP-MRI) dataset van elke individuele patiënt.

Twee MP-MRI datasets waren beschikbaar om de studies in deze thesis uit te voeren, opgemeten in de universitaire ziekenhuizen van Leuven en Gent. Ze bevatten beiden cMRI, PWI, DWI en MRSI, maar hun scanprotocols verschillen. Terwijl voor tumor segmentatie meestal enkel cMRI wordt beschouwd, worden in deze thesis alle beschikbare MP-MRI data gecombineerd, en de toegevoegde waarde van de individuele MRI modaliteiten wordt nagegaan.

Een hiërarchische variant van NMF (hNMF) wordt beschouwd, die in wezen bestaat uit 2 NMF niveaus. hNMF is in staat om een geldige weefsel differentiatie te bekomen, alsook een competitieve segmentatie in termen van de Dice score. Voor hoog-gradige gliomen blijken de Dice scores voor de volledige MP-MRI dataset significant hoger te zijn dan voor gereduceerde MP-MRI datasets, waarbij telkens één MRI modaliteit wordt weggelaten. De computationele efficiëntie van het hNMF algoritme wordt verbeterd door middel van geavanceerde initialisatie methodes. In het bijzonder wordt het successieve projectie algoritme (SPA) beschouwd. Hoewel SPA voornamelijk gebruikt wordt als tool om de NMF bronnen rechtstreeks te bepalen, blijkt dat men betere segmentatie bekomt door SPA als initialisatiemethode toe te passen. hNMF wordt ook vergeleken met andere vaak gebruikte ongesuperviseerde classificatietechnieken, waaronder enkele single-level NMF en clustering methodes. hNMF overtreft de andere methodes op basis van de MP-MRI datasets van beide ziekenhuizen. Dankzij zijn hiërarchische structuur kan hNMF weefseltypes met gelijkaardige signatuur beter onderscheiden, zoals niet-contrastcapterende tumor en oedeem.

Ongesuperviseerde segmentatie algoritmes moeten voorkennis integreren om competitief te kunnen zijn. Een semi-gesuperviseerde NMF segmentatie techniek wordt beschouwd die gebruik maakt van voxel selectie om de pathologische NMF bronnen te initialiseren. L1-regularisatie wordt gebruikt om spatiële consistentie en schaarste van de weefsel proporties te bevorderen. Een morfologische post-processing maakt gebruik van de locatie van de geselecteerde voxels en connectiviteitsbeperkingen tussen de weefsels om de segmentatie verder te verbeteren. Zorgvuldige voxel selectie in de deelgebieden van de tumor blijkt noodzakelijk om robuuste segmentatie te bekomen. Semi-automatische NMF segmentatie werd ook toegepast op de publiekelijk beschikbare BRATS 2013 Leaderboard dataset. Competitieve segmentatie resultaten werden bekomen, in het bijzonder voor actieve tumor, waarvoor onze methode de beste score haalt van alle deelnemers. Deze thesis toont aan dat ongesuperviseerde segmentatie algoritmes kunnen concurreren met state-of-the-art gesuperviseerde methodes mits gebruik te maken van toepassingsgerichte voorkennis.

Abbreviations

ADC	Apparent diffusion coefficient.
aHALS	Accelerated hierarchical alternating least squares.
AIF	Arterial input function.
ALS	Alternating least squares.
BRATS	Multi-modal brain tumor segmentation challenge.
CBF	Cerebral blood flow.
CBV	Cerebral blood volume.
Cho	Choline.
cMRI	Conventional magnetic resonance imaging.
CNS	Central nervous system.
Cre	Creatine.
CSF	Cerebro-spinal fluid.
CT	Computed tomography.
DKI	Diffusion kurtosis imaging.
DSC-MRI	Dynamic susceptibility-weighted magnetic resonance imaging.
DTI	Diffusion tensor imaging.
DWI	Diffusion-weighted imaging.
EPI	Echo-planar imaging.
FA	Fractional anisotropy.

FCM	Fuzzy C-means clustering.
FLAIR	T_2-weighted imaging with fluid- attenuated inversion recovery.
FOV	Field of view.
GBM	Glioblastoma.
Gln	Glutamine.
Glu	Glutamate.
Glx	Glutamine+glutamate.
Gly	Glycine.
GMM	Gaussian mixture modelling.
HALS	Hierarchical alternating least squares.
HGG	High-grade glioma.
HMRF	Hidden Markov random fields.
hNMF	Hierarchical Non-negative matrix factorization.
Lac	Lactate.
LGG	Low-grade glioma.
Lip1	Lipids at 1.3ppm.
Lip2	Lipids at 0.9ppm.
MD	Mean diffusivity.
mI	Myo-inositol.
MICCAI	International conference on Medical Image Computing and Computer Assisted Interventions.
MK	Mean kurtosis.
MP-MRI	Multi-parametric MRI.
MPFIR	Maximum-phase finite impulse response.
MRI	Magnetic resonance imaging.
MRS	Magnetic resonance spectroscopy.
MRSI	Magnetic resonance spectroscopic imaging.
MTT	Mean transit time.
NAA	N-acetyl aspartate.
NMF	Non-negative matrix factorization.
NMR	Nuclear magnetic resonance.
NNDSVD	Non-negative double singular value decomposition.

NNLS	Non-negative least squares.
PPV	Positive predictive value.
PWI	Perfusion-weighted imaging.
RANO	Response assessment in neuro-oncology.
rCBV	Relative cerebral blood volume.
RF	Radio frequency.
ROI	Region of interest.
rRandom	Repetitive random initialization.
SC	Spectral clustering.
SDF	Structured data fusion.
SNR	Signal-to-noise ratio.
SPA	Successive projection algorithm.
T1c	T_1-weighted imaging with contrast enhancement.
TE	Echo time.
TI	Inversion time.
TR	Repetition time.
UZ Gent	University hospital of Gent.
UZ Leuven	University hospitals of Leuven.
VOI	Volume of interest.

Contents

Abstract iii

Contents xi

List of Figures xvii

List of Tables xxiii

1 Introduction 1

1.1 Brain tumors 1

1.1.1 Epidemiology 1

1.1.2 Gliomas 2

1.2 Magnetic resonance imaging 7

1.2.1 Principles of MRI 7

1.2.2 Conventional MRI 10

1.2.3 Perfusion-weighted imaging 14

1.2.4 Diffusion-weighted imaging 20

1.2.5 Magnetic resonance spectroscopic imaging 24

1.3 Brain tumor segmentation 31

1.3.1 Segmentation in treatment planning and follow-up . . . 31

1.3.2 Segmentation methods 32
1.3.3 Segmentation using advanced MRI modalities 34
1.4 Aims of the thesis 35
1.5 Outline of the thesis and personal contributions 37
1.6 Conclusion 39

2 Non-negative matrix factorization and validation metrics 41
2.1 Non-negative matrix factorization 41
2.1.1 NMF model 41
2.1.2 NMF algorithms 42
2.2 Validation metrics 45
2.2.1 Dice score 45
2.2.2 Positive predictive value 46
2.2.3 Sensitivity 46
2.2.4 Cosine similarity 46
2.3 Conclusion 47

3 Multi-parametric MRI datasets and pre-processing 49
3.1 UZ Leuven dataset 50
3.1.1 Patient population 50
3.1.2 cMRI 50
3.1.3 PWI 51
3.1.4 DKI 52
3.1.5 MRSI 53
3.1.6 MP-MRI dataset 57
3.2 UZ Gent dataset 58
3.2.1 Patient population 58
3.2.2 cMRI 58

3.2.3 PWI 59
3.2.4 DWI 59
3.2.5 MRSI 60
3.2.6 MP-MRI dataset 61

4 Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using MP-MRI data 63

4.1 Introduction 64
4.2 MP-MRI dataset 65
4.3 Hierarchical non-negative matrix factorization 67
4.4 Results 72
4.4.1 Dice scores 75
4.4.2 Correlation coefficients 77
4.4.3 Influence of the averaged features 79
4.5 Discussion 80
4.5.1 Comparison with previous NMF studies 80
4.5.2 Reduced MP-MRI datasets and cMRI only 82
4.5.3 Tumor segmentation 84
4.5.4 Unsupervised vs supervised classification 85
4.5.5 Voxel-wise vs ROI analysis 85
4.5.6 Conclusion 86

5 The successive projection algorithm as an initialization method for brain tumor segmentation using NMF 87

5.1 Introduction 88
5.2 MP-MRI dataset 90
5.3 NMF and initialization methods 91
5.3.1 NMF methods 91

5.3.2 Geometrical interpretation of NMF 93
5.3.3 Initialization methods 94
5.3.4 Validation 96
5.4 Results 97
5.5 Discussion 100
5.5.1 NMF performance 100
5.5.2 Convergence and computational cost 101
5.5.3 SPA: initialization or direct source extraction 102
5.6 Conclusion 102

6 Comparison of unsupervised classification methods for brain tumor segmentation using MP-MRI data 105

6.1 Introduction 105
6.2 MP-MRI datasets 108
6.2.1 UZ Gent 108
6.2.2 UZ Leuven 109
6.2.3 Image co-registration and voxel selection 109
6.3 Methods 110
6.3.1 Non-negative matrix factorization 110
6.3.2 Clustering 113
6.3.3 Initialization 115
6.3.4 Validation 116
6.4 Results 116
6.4.1 UZ Gent 116
6.4.2 UZ Leuven 119
6.4.3 Computational cost 120
6.4.4 Data distribution 121
6.5 Discussion 122

6.5.1 NMF vs clustering 122
6.5.2 hNMF performance 123
6.5.3 Spatial regularization 124
6.5.4 UZ Gent and UZ Leuven datasets 124
6.5.5 Initialization methods 125
6.5.6 Supervised vs unsupervised classification 126
6.6 Conclusion 127

7 Semi-automated brain tumor segmentation on MP-MRI data using regularized NMF 129

7.1 Introduction 130
7.2 MP-MRI dataset 132
7.3 Methodology 133
7.3.1 Regularized NMF 133
7.3.2 NMF initialization 134
7.3.3 Morphological post-processing 135
7.3.4 Validation 136
7.4 Results 137
7.5 Discussion 141
7.6 Conclusion 144

8 Application of semi-automated regularized NMF to the BRATS 2013 Leaderboard dataset 147

8.1 Introduction 148
8.2 BRATS 2013 Leaderboard dataset 149
8.3 Methodology 149
8.3.1 Bias field correction 149
8.3.2 NMF-based semi-automated segmentation 150

8.3.3 Validation 151
8.4 Results 152
8.5 Discussion 155
8.6 Conclusion 159

9 Conclusions and future perspectives **161**
9.1 Conclusions 162
9.2 Future perspectives 166

A Validation scores individual UZ Leuven patients **171**

Bibliography **175**

Curriculum vitae **207**

List of publications **209**

List of Figures

1.1 Average number of newly diagnosed brain and CNS tumors per year and age-specific incidence rates per 100,000 in the UK. *Source: cruk.org/cancerstats.* 2

1.2 Distribution of gliomas by histology subtypes. *Source: CBTRUS Statistical Report: NPCR and SEER, 2007–2011 [184].* 4

1.3 A collection of precessing spins around an external magnetic field B_0 results in a net magnetization vector M in the direction of B_0. 8

1.4 A 90° RF-pulse causes the magnetization vector M to be rotated about the B_1-axis into the transversal plane. 9

1.5 cMRI examination of a grade II astrocytoma patient. Left to right: T1+contrast, T2, FLAIR, manual delineation of the tumor. 12

1.6 cMRI examination of a grade IV GBM patient. Left to right: T1+contrast, T2, FLAIR, manual delineation of active tumor (red), necrosis (green) and edema (blue). 12

1.7 A: Signal intensity decrease during passage of contrast agent bolus is measured using DSC-MRI. B: Change in the relaxation rate (ΔR_2^*) is calculated from signal intensity, and a baseline subtraction method is applied to measured data. C: Corrected ΔR_2^* curve. D: rCBV is proportional to the area under curve. *Source: Cha et al., Radiology, 2002 [40].* 17

1.8 Diffusion-sensitized spin echo sequence. The spin echo sequence is modified by the addition of two gradient pulses (grey blocks) which induce and then reverse a spatially-dependent phase shift. *Source: Winston et al., Quant Imaging Med Surg, 2012 [278].* . . 21

1.9 Isotropic unrestricted and restricted diffusion and anisotropic diffusion. *Source: Mukherjee et al., Am Soc Neuroradiology, 2008 [173].* 22

1.10 Short-TE ^{1}H-MRS spectrum acquired within a GBM. The main metabolites of interest are indicated, right to left: lipids at 0.9ppm (Lip2), lactate (Lac), lipids at 1.3ppm (Lip1), N-acetyl-aspartate (NAA), glutamate+glutamine (Glx), creatine (Cre), choline (Cho), glycine (Gly) and myo-inositol (mI)) 26

1.11 T_1-weighted structural image overlaid with the MRSI grid. Detailed spectra are shown for a normal-appearing white matter voxel (A) and for a cortical gray matter voxel (B). *Source: Chard et al., Brain, 2002 [44].* 27

1.12 Left: T_2-weighted image of a GBM patient overlaid with the MRSI grid. Right: quantified metabolite maps obtained from short-TE ^{1}H-MRSI acquisition. (a)myo-inositol, (b)choline, (c)creatine, (d)glutamine+glutamate, (e)N-acetyl aspartate, (f)lactate, (g)lipids at 1.3ppm, (h)lipids at 0.9ppm. 28

1.13 Schematic overview of the thesis. 38

3.1 Cre maps obtained using QUEST and AQSES for a GBM patient with low-quality spectral data. The MRSI region of interest is marked in green on T1c on the left. 55

3.2 Combined (A) and separate (B) quantification of mI and Gly in a GBM patient. Increased Gly and reduced mI (B) are common in GBM. 56

3.3 Co-registered set of MP-MRI maps of a GBM patient from UZ Leuven. 'Seg' indicates manual segmentation of active tumor (red), necrosis (green) and edema (blue). Color scaling differs among the different metabolites. 57

3.4 Co-registered set of MP-MRI maps of a GBM patient from the UZ Gent dataset. Color scaling differs among the different metabolites. 61

4.1 Left: abundances maps of the 2 sources after hNMF step1. Voxel selection for the second level of NMF is shown for t=0.5, t=0.7 and t=1. 69

4.2 Schematic overview of the 3 steps of the hNMF procedure. . . . 70

4.3 MP-MRI maps of different modalities and hNMF results for GBM patient 16. The shown input maps are: T2 (A), T1c (B), MD (C), MK (D), rCBV (E), Cho (F), Cre (G) and Lip (H). The green box indicates the MRSI ROI. hNMF abundance maps are shown for tumor (I), necrosis (J) and edema (K). The residual error map is shown in (L) and the pathological hNMF sources in (M). 73

4.4 Comparison of the hNMF results for the different MRI datasets for GBM patient 16. hNMF abundance maps are shown for active tumor and necrosis. A comparison of the segmentation obtained with hNMF (blue) and by the radiologist (green) is shown for active tumor and for the tumor core. Cyan indicates segmentation overlap. 74

5.1 Geometrical representation of exact NMF for a rank-2 non-negative input matrix $X \in R_+^{3\times 10}$. 93

5.2 Co-registered set of MRI images, showing T1c(A), rCBV(B), MD(C) and Lac(D). The green frame in the images of the upper row marks the MRSI region of interest. Segmentation results are shown for active tumor and necrosis, respectively, for aHALS NMF(E,F), Convex NMF(G,H), SDF NMF(I,J) and hNMF(K,L) with SPA initialization. NMF segmentation is shown in blue, manual segmentation by a radiologist in green and segmentation overlap in cyan. 97

5.3 Convergence plots for aHALS NMF (A), Convex NMF (B) and SDF NMF (C) with the different initialization methods. The residual error, $\|X - WH\|_F$, is shown on a log scale. For rRandom and FCM, the shown curve corresponds to the selected run with the lowest residual error. 99

6.1 Schematic overview of the adapted hNMF algorithm. 112

6.2 a) Some of the MP-MRI images of a grade III oligo-astrocytoma patient from the UZ Gent dataset. 'Seg' indicates manual segmentation by a radiologist (active tumor=red, edema=blue). The ROI is delineated in green. b) Segmentation results (using kmeans++ initialization) of the NMF methods (top row) and the clustering methods (bottom row). 118

6.3 Active tumor and necrosis datapoints projected onto the plane formed by the first and second principal component for an UZ Leuven GBM patient (left) and for an UZ Gent GBM patient (right). 122

7.1 Illustration of morphological post-processing after initial semi-automated NMF based segmentation of necrosis (A) and active tumor (C). Step 1: false positives are removed by withholding only the connected components closest to the user-defined seeding points (marked by an 'x' in A and C). Step 2: spatial adjacency of the connected components in the preliminary necrosis mask (green) to the withheld active tumor mask (red) is verified in E. The final necrosis mask is shown in yellow in F. 136

7.2 First row: coregistered MP-MRI maps of a GBM (patient 3), left to right: T1c, FLAIR, rCBV, MD. Second row, left to right: NMF abundance maps for active tumor, necrosis and edema. The final segmentation masks are shown on the right for active tumor (red), necrosis (yellow) and edema (blue). 137

7.3 Boxplots showing the dispersion of the Dice scores for active tumor. Boxplots show quartile ranges of the Dice scores, '+' indicates outliers. 138

7.4 Boxplots showing the dispersion of the Dice scores for the tumor core. Boxplots show quartile ranges of the Dice scores, '+' indicates outliers. 139

7.5 Boxplots showing the dispersion of the Dice scores for the whole tumor region. Boxplots show quartile ranges of the Dice scores, '+' indicates outliers. 139

7.6 A close-up of a GBM lesion on an axial T1c slice (A). The manually segmented necrotic region is shown in yellow (B), the selected necrotic voxel is marked in red. Segmentation of the active tumor region (C) and necrotic region (D) on several slices, blue indicates NMF segmentation, green indicates manual segmentation and cyan indicates overlap. 145

8.1 Uncorrected (left) and bias field corrected (right) axial T1c slice of HGG patient 131. 150

8.2 Left to right: Axial T1c slice of HGG patient 205, initial segmentation of all tissue sources after semi-automated NMF, and the final segmentation of the tumor subcompartments after morphological post-processing (red=enhancing tumor, blue=necrosis and green=edema). 155

List of Tables

4.1 Schematic overview of the MP-MRI protocols and derived parameters of the UZ Leuven dataset. 66

4.2 Mean, standard deviation and range of the Dice scores for the HGG patients. Results are reported for active tumor, the tumor core and the whole tumor region using the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only. p-values report statistical significance of higher Dice scores for the full MP-MRI dataset (* indicates statistical significance for $p<0.05$). 76

4.3 Mean, standard deviation and range of the Dice scores for the LGG patients. Results are reported for the whole tumor region using the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only. p-values report statistical significance of higher Dice scores for the full MP-MRI dataset (* indicates statistical significance for $p<0.05$). 77

4.4 Mean, standard deviation and range of the correlation coefficients for the HGG patients. Results are reported for active tumor, necrosis and edema using the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI and for cMRI data only. p-values report statistical significance of higher correlation coefficients for the full MP-MRI dataset (* indicates statistical significance for $p<0.05$). 78

4.5 Mean, standard deviation and range of the correlation coefficients for the LGG patients. Results are reported for active tumor using the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI and for cMRI data only. p-values report statistical significance of higher correlation coefficients for the full MP-MRI dataset (* indicates statistical significance for p<0.05). 79

4.6 Comparison of the mean Dice scores and correlation coefficients for hNMF with and without averaged non-metabolic features. Results are reported for HGGs and LGGs separately. * indicates statistical significance of higher values, using the one-tailed Wilcoxon signed rank test with p<0.05. 80

5.1 Schematic overview of the MP-MRI protocols and derived parameters of the UZ Gent dataset. 90

5.2 Comparison of the mean Dice scores and their standard deviation between different initialization methods.* indicates statistically significantly higher Dice scores with SPA initialization compared to direct SPA endmember extraction (right column), using a one-tailed Wilcoxon signed rank test ($p < 0.05$). 98

5.3 Mean number of iterations to reach convergence for the different initialization methods. Convergence tolerance was set to 10^{-5} and the maximum number of iterations to 10000. 99

6.1 Schematic overview of the MP-MRI protocols and derived parameters of the 2 datasets acquired at UZ Gent and UZ Leuven. 108

6.2 Segmentation results for the UZ Gent dataset when using kmeans++ initialization. Mean Dice score ± standard deviation is reported for active tumor, necrosis, edema, the tumor core and the whole tumor. The number of undetected cases is reported for active tumor, necrosis and edema. 117

6.3 Segmentation results for the UZ Gent dataset when using SPA initialization. Mean Dice scores ± standard deviation are reported for active tumor, necrosis, edema, the tumor core and the whole tumor. The number of undetected cases is reported for active tumor, necrosis and edema. 117

6.4 Segmentation results for the UZ Leuven dataset when using kmeans++ initialization. Mean Dice scores ± standard deviation are reported for active tumor, necrosis, edema, the tumor core and the whole tumor. The number of undetected cases is reported for active tumor, necrosis and edema. 119

6.5 Segmentation results for the UZ Leuven dataset when using SPA initialization. Mean Dice scores ± standard deviation are reported for active tumor, necrosis, edema, the tumor core and the whole tumor. The number of undetected cases is reported for active tumor, necrosis and edema. 120

6.6 Average computation time (in seconds) per patient on the UZ Gent dataset using kmeans++ initialization. 121

7.1 Schematic overview of the MP-MRI protocols and derived parameters of the UZ Gent dataset. 133

7.2 Mean Dice scores [%] for NMF without regularization, with spatial regularization (λ=0.1), with spatial regularization and sparseness (λ=0.1), and using only cMRI data with spatial regularization and sparseness. For each tissue class, the first column reports the mean of the 25th percentile Dice score across all patients, the second column the mean of the mean Dice score across all patients and the third column the mean of the 75th percentile Dice score. 140

8.1 Mean Dice score, PPV and sensitivity values of the enhancing tumor, tumor core and whole tumor region on the BRATS 2013 Leaderboard dataset. The first 5 rows rank the best performing groups that participated in the BRATS 2013 challenge. Our results are reported on row 6 (Sauwen) and rows 7 and 8 show other results from literature. 153

8.2 Dice score ranges (min-max) for the individual patients over 3 repeated runs using different voxel selection. Color coding: green = range<5%, yellow = range<10%, orange = range<15% and red = range>15%. 154

A.1 Dice scores for active tumor, the tumor core and the whole tumor region for the UZ Leuven HGG patients. Results are reported for the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only. 172

A.2 Correlation coefficients for active tumor, necrosis and edema for the UZ Leuven HGG patients. Results are reported for the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only. 173

A.3 Dice scores for the whole tumor region of the UZ Leuven LGG patients. Results are reported for the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only. 174

A.4 Correlation coefficients for the active tumor region of the UZ Leuven LGG patients. Results are reported for the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only. 174

Chapter 1

Introduction

This chapter introduces some of the main topics that will be discussed throughout the thesis. The first section focuses on brain tumors and in particular on gliomas, as they are the most common type of primary brain tumors. Challenges and limitations regarding the diagnosis and treatment of gliomas are explained. Magnetic resonance imaging (MRI) has become the imaging modality of choice in managing glioma patients. The principles of MRI and the different MRI modalities are discussed, including their added value in the clinical management of gliomas. Brain tumor segmentation is a crucial step in treatment planning and during follow-up. The current status of brain tumor segmentation in clinical practice, as well as the main advancements in automated segmentation are discussed. The main goals of the thesis are formulated, and a structured outline of the thesis is given in the final section.

1.1 Brain tumors

1.1.1 Epidemiology

Brain tumors originate from an abnormal and uncontrolled cell growth inside the cranium. With an incidence rate of about 25 new cases per year per 100,000 [124], brain tumors are not among the most common tumors in the western population. But they are among the most fatal cancers [59]. It is estimated that 77,670 persons will be diagnosed with a primary brain tumor in the United States in 2016 [184]. About one-third of these tumors are expected to be malignant, while the rest are either benign or borderline-malignant. An estimated 16,616

deaths will be attributed to primary malignant brain and central nervous system tumors in the US in 2016. The five-year survival rate following diagnosis of a primary malignant brain tumor is 34.4% [184]. Incidence of brain tumors is related to age, with the highest incidence rates overall being in older males and females (see Figure 1.1). In contrast to most cancer types, brain tumors also occur relatively frequently at younger ages. Age-specific incidence rates remain relatively stable from infancy to around age 25-29, before increasing more sharply with the highest rates in the 90+ age group.

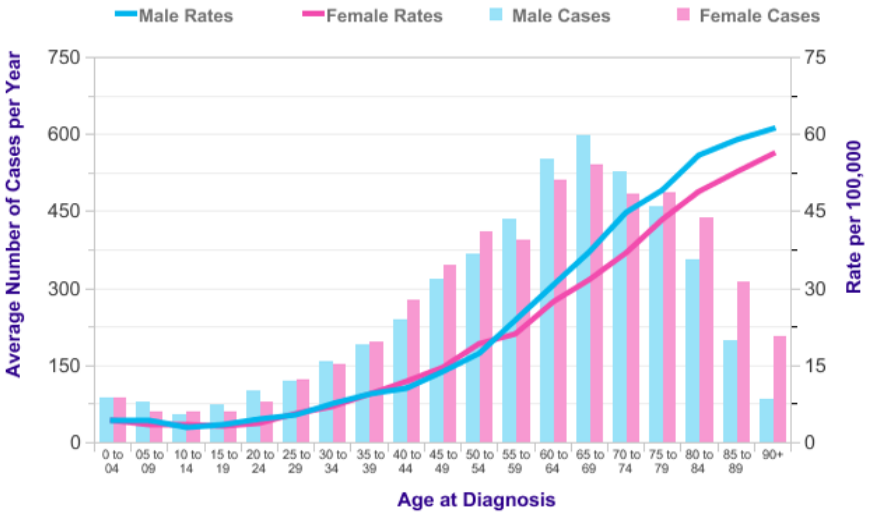

Figure 1.1: Average number of newly diagnosed brain and CNS tumors per year and age-specific incidence rates per 100,000 in the UK. *Source: cruk.org/cancerstats.*

1.1.2 Gliomas

Classification

Gliomas are the most common type of primary brain tumors. They represent about 30% of all primary brain tumors and 80% of all malignant primary brain tumors [184]. The majority of these tumors occur in the frontal, temporal, parietal, and occipital lobes. The term glioma refers to tumors that have histologic features similar to glial cells: astrocytomas microscopically show most resemblance to astrocytes, oligodendrogliomas have most resemblance

to oligodendrocytes and ependymomas histologically mimic ependymal cells. Oligoastrocytomas present with an appearance of mixed glial cell origin. According to the World Health Organization (WHO) classification system, gliomas are further categorized based on a grading scheme that represents a malignancy scale [157]. Grading is based primarily on histologic criteria, such as cell density, nuclear and cellular atypia, number of mitoses, and vascular endothelial proliferation. WHO classifies gliomas into 4 grades, with grades I and II referred to as low-grade gliomas (LGGs) and grades III and IV termed high-grade gliomas (HGGs). Grades I and II may be considered as semi-malignant tumors that come with a better prognosis, whereas grade III and IV tumors are malignant tumors that almost certainly lead to a subject's death.

Grade I gliomas are also termed pilocytic astrocytomas. Pilocytic astrocytomas mostly appear in children and in young adults (age 0 to 20 years). These tumors are usually benign and slow-growing. Under the microscope, they are seen to be composed of bipolar cells with long hairlike processes. These tumors are usually well-circumscribed and total surgical resection is often possible. The 5-year survival rate for pilocytic astrocytoma is 94% [184].

Grade II gliomas show increased cellularity and nuclear atypia. Grade II astrocytomas are diffuse, infiltrating the surrounding healthy brain and often progressing towards more malignant forms over the course of years. The 5-year survival rate for grade II astrocytoma is 47%. Grade II oligodendrogliomas and ependymomas have a better overall prognosis.

Grade III gliomas are also termed anaplastic gliomas. They exhibit higher cellularity, more pronounced infiltrative growth as well as increased cell proliferation (i.e. fast divisive cell growth). Grade III astrocytomas don't demonstrate microvascular proliferation (i.e. formation of new microvessels), whereas grade III oligodendrogliomas and oligoastrocytomas do show some degree of neo-vascularization. The 5-year survival rate for anaplastic astrocytoma is 27%.

Grade IV gliomas are often referred to as glioblastoma (GBM). GBM is not only the most malignant, but also the most common type of glioma (see Figure 1.2). These tumors are highly vascular and they demonstrate abundant infiltration. Presence of necrosis, i.e. regions of dead cell tissue at the centre of the tumor, is the main hallmark for differentiating GBMs from grade III gliomas. The active part of the tumor consists of highly dense regions of anaplastic tumor cells. GBMs have a dramatically low 5-year survival rate of 5%.

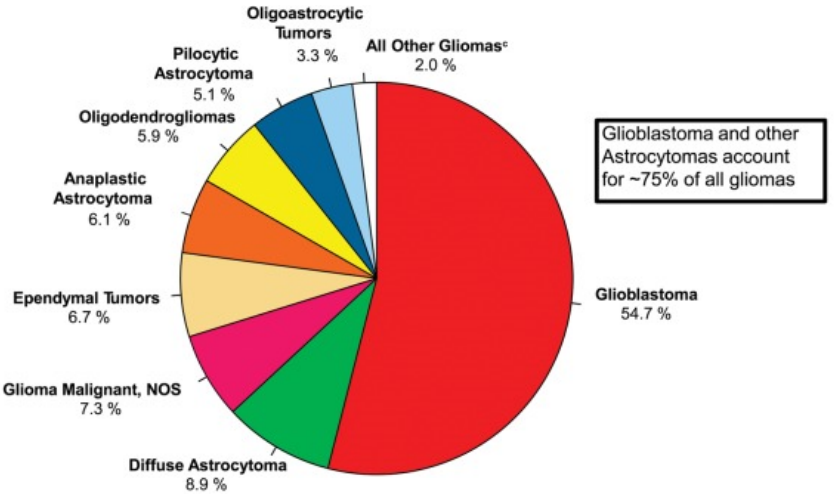

Figure 1.2: Distribution of gliomas by histology subtypes. *Source: CBTRUS Statistical Report: NPCR and SEER, 2007–2011 [184].*

Diagnosis

The most common symptoms of glioma patients are related to increased intracranial pressure caused by the expanding tumor: seizures, headache and weakness in the arms, face or legs. Seizures are more prevalent in LGGs than in HGGs [172]. Headache is the worst symptom in approximately 50% of patients. Features suggestive of a brain tumor in a patient complaining of headache include nausea and vomiting [75]. Other common symptoms include fatigue, sleep disturbance and cognitive impairment [235].

Upon suspicion of a brain tumor, initial diagnosis relies on the medical history and age of the patient, a neurological examination and neuro-imaging. Magnetic resonance imaging (MRI) has become the imaging modality of choice for the management of brain tumor patients. MRI shows excellent soft tissue contrast and different structures of interest can be enhanced by varying excitation and repetition times. Differential diagnosis is based on conventional MRI (cMRI) and relies on radiological features, like the size and location of the lesion, disruption of the blood-brain barrier and the presence of necrosis.

Although clinical and neuro-imaging features can be highly suggestive, the gold standard for diagnosis and classification of gliomas is based on histologic examination of pathological tissue samples. Samples are obtained either through

a biopsy or upon surgical removal of the gross tumor volume. Whereas no variable predicts prognosis more precisely, histologic classification comes with some limitations. First of all, visual assessment of histopathological specimens is subject to inter- and intra-observer variability [254]. A second limitation is related to the high heterogeneity that is known to occur in HGGs. Different cellular phenotypes and malignancy grades can occur within the same lesion [190]. Due to this heterogeneity, biopsy sampling might not represent the tissue characteristics of the whole tumor, and there is a risk of underdiagnosing if the biopsy is not taken from the most malignant part of the tumor. Underdiagnosing an HGG would have the implication of depriving patients from the radical treatment options they need.

Treatment

Treatment of gliomas is customized to the individual patient and may include surgery, radiation therapy and/or chemotherapy. Surgery is the most common initial treatment. Maximal safe resection is performed, meaning that the gross tumor volume is maximally being removed, while leaving the surrounding healthy brain intact as much as possible. The extent of tumor resection has been shown to be an important prognostic factor for patient survival time in GBMs [130]. Partial resection is performed when tumors are located within or adjacent to eloquent brain regions, and no surgery is performed when a tumor is located in a critical region of the brain. The benefits of resecting as much tumor as possible to increase survival must be carefully balanced with the risks of compromising neurologic function and decreasing quality of life. A unique feature of both LGGs as well as HGGs is that they tend to infiltrate the surrounding healthy brain. cMRI tends to underestimate the full tumor extent, as it is not sensitive to infiltrated regions of peritumoral edema or healthy brain [39]. The diffuse growth pattern of gliomas precludes complete tumor removal, and is one of the main reasons of therapeutic failure [52].

Radiation therapy is the main non-surgical treatment for gliomas. It is well established that radiation therapy delays tumor recurrence and prolongs life in HGGs [265], however nearly all patients still develop recurrent tumor within the treated volume [42]. During radiotherapy planning, the tumor is delineated based on cMRI to define the target volume.

Chemotherapy may also be delivered as adjuvant therapy, possibly in combination with radiation therapy. The success of chemotherapy is hampered by the marked heterogeneity within gliomas [209]. Especially in areas where the original tissue structure is relatively preserved, the blood-brain barrier may form an obstacle for optimal delivery of chemotherapeutics to infiltrative tumor

cells. Recently, temozolomide treatment in GBM patients has been shown to result in modest improvement of median overall survival and 2-year survival rate [244].

Pilocytic astrocytomas are not infiltrative and they have a favorable diagnosis. Curable treatment is often possible by complete surgical resection of the lesion. Even when the resection is incomplete, radio- and chemotherapy are only considered when there is evidence of tumor growth. The role of adjuvant radio- and/or chemotherapy in grade II astrocytomas is uncertain, but is often postponed until progression is observed. Radiation therapy is widely used to treat residual disease after partial resection in diffuse astrocytomas. Anaplastic astrocytomas are usually treated with surgery, radiation therapy and adjuvant chemotherapy. These tumors are generally more responsive to chemotherapy than GBMs. Maximal safe resection in combination with radiotherapy is also recommended in diffuse and anaplastic oligodendrogliomas. Such tumors with deletion of the chromosome 1p/19q have been reported to be highly responsive to chemotherapy using alkylating agents [33]. Standard treatment of GBM consists of maximal safe resection, fractionated radiation therapy and chemotherapy with temozolomide. It has been reported that GBM patients treated with adjuvant chemotherapy and radiation therapy had significant survival benefit compared to patients treated with radiation therapy alone [244].

Despite considerable advancements in our understanding of the biology of gliomas, the prognosis for these tumors remains appalling with conventional treatment. The increasing knowledge on molecular pathways and genetic deficits involved in glioma growth, opens up new perspectives in developing innovative treatment strategies. This has led to the recent development of a number of targeted therapies, i.e. therapies that focus on the tumor cells without having major effects on the surrounding healthy brain tissue. Anti-angiogenic therapies comprise the most widely studied new group of treatments. They act upon the vascularization of malignant gliomas, either by normalizing the tumor vasculature or by inhibiting vascular ingrowth. The most widely studied agent is bevacizumab [77]. This agent normalizes the tumor vasculature, which may improve the delivery of chemotherapeutic drugs. Another emerging strategy involves the stimulation of an anti-tumor immune response. Immune therapy is appealing because it would induce an immune response specifically targeting tumor cells without the risk of damaging normal structure. Dendritic cell therapy is one of the most common approaches, in which tumor cell antigens are merged with dendritic cells [255]. Other recently emerging therapies include high-intensity focused ultrasound [17], photo-dynamic therapies [15] and engineering of nanoparticles for drug delivery or thermal treatment [47]. Some of these new therapies are already being considered in clinical trials, whereas others are still at the experimental stage. Further research is also required to investigate if

and how these new therapies should be combined with conventional treatment options.

1.2 Magnetic resonance imaging

1.2.1 Principles of MRI

One would have to dig into quantum mechanics to thoroughly understand the principles of MRI. In the context of this thesis, such an explanation would lead us to far, as it is not the subject of research. However, all aspects of basic MRI can alternatively be explained using intuitive classical mechanics, which is considered in the text below.

MRI visualizes a quantum-physical property of tissue called the nuclear spin. Making this magnetic property of nuclear spins visible requires the external application of 2 magnetic fields: a strong static magnetic field B_0 and an oscillating magnetic field B_1 in a plane perpendicular to B_0. When placed in an external magnetic field B_0, the individual spins are aligned with B_0. Individual spins are either parallel to B_0, corresponding to a low energy spin state, or anti-parallel to B_0, corresponding to a less stable higher energy spin state. More spins will occupy the lower energy state, resulting in a net magnetization vector M along the B_0 axis (see Figure 1.3). The spins precess with a specific frequency f_0 which is proportional to B_0 [151]. At this stage, the precession of the net magnetization vector can not be detected, as its rotational axis coincides with the vector itself.

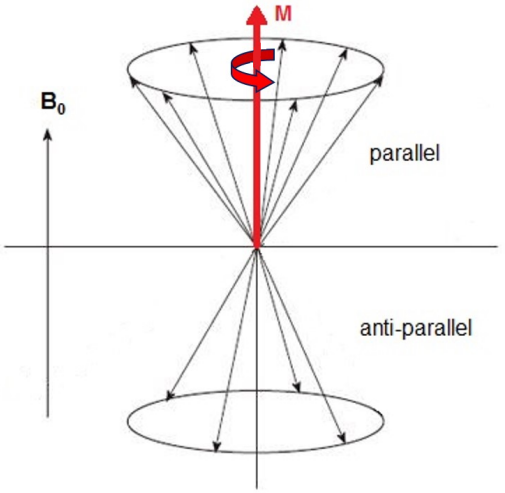

Figure 1.3: A collection of precessing spins around an external magnetic field B_0 results in a net magnetization vector M in the direction of B_0.

To induce transitions between both energy states, an oscillating magnetic field B_1 is applied in a plane perpendicular to B_0. If the frequency of B_1 corresponds to the energy difference between both states, nuclei at the lower energy state can absorb the radio frequency (RF-)energy and flip to the high-energy state. Nuclei undergoing this transition are said to be in magnetic resonance. Another consequence of applying B_1 is that all individual nuclear spins become phase coherent. While the net magnetization vector M only has a longitudinal component (i.e. along the B_0-axis) at equilibrium, excitation with an RF-pulse through B_1 introduces a transversal perpendicular component (see Figure 1.4). As M precesses around the longitudinal axis, its transversal component introduces an oscillating magnetic flux which yields a changing voltage in receiver coils to give the RF output signal.

When the oscillating magnetic field is turned off, spins return to their equilibrium state in a process which is called relaxation. Relaxation consists of several aspects:

- T_1- or spin-lattice relaxation: refers to the recovery of the longitudinal magnetization component, which occurs at an exponential rate with a time constant T_1.
- T_2- or spin-spin relaxation: describes the return to equilibrium of the

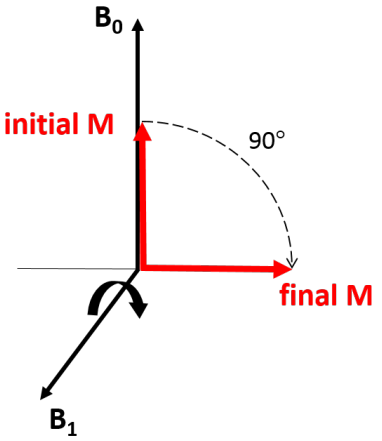

Figure 1.4: A 90° RF-pulse causes the magnetization vector M to be rotated about the B_1-axis into the transversal plane.

transverse magnetization component, which occurs at an exponential rate with a time constant T_2.

- T_2^*-relaxation: describes the actual decay of the transverse magnetization signal due to local magnetic field inhomogeneities. The exponential decay is defined by the time constant T_2^*.

Hydrogen atoms possess a nuclear spin and they are abundantly present in the human body, particularly in water and fat. For this reason, conventional MRI techniques essentially map the location of water and fat in the body. Tissue contrast is obtained by differences in relative hydrogen abundance and by the fact that hydrogen atoms in different tissue types exhibit different relaxation times. MRI is a highly versatile imaging technique, as it can produce different contrast between tissue types by varying the parameters of the pulse sequence:

- the echo time (TE) is the time between the RF excitation pulse and the peak in the MR signal induced in the receiver coil, at which data acquisition starts.
- the repetition time (TR) is the time between successive pulse sequences. It determines how much the longitudinal magnetization can recover after each pulse.

The design of dedicated pulse sequences and variation of the sequence parameters have led to the wide versatility in imaging capabilities. It allows MRI to deliver structural as well as functional imaging information. The next paragraphs will discuss some of the most common MRI techniques in detail.

1.2.2 Conventional MRI

In neuro-imaging, cMRI has largely replaced computed tomography (CT) as the imaging modality of choice. cMRI provides better spatial resolution in combination with excellent soft tissue contrast. Furthermore, it does not expose the patient to potentially harmful ionizing radiation. MRI makes it possible to produce markedly different types of tissue contrast by varying excitation and repetition times.

T_1-weighted imaging

T_1-weighted imaging is one of the basic pulse sequences in MRI and demonstrates differences in T_1 relaxation time among different tissue types. It relies upon the longitudinal relaxation of a tissue's magnetisation vector. T_1-weighted images are obtained using a short TE and short TR. Not all tissues get back to equilibrium equally fast, and a tissue's T_1 reflects the amount of time its protons' spins realign with the main magnetic field. Fat quickly realigns its longitudinal magnetization with B_0, and it therefore appears bright (hyper-intense) on a T1 weighted image. Water has much slower longitudinal magnetization realignment, and therefore has less transverse magnetization after an RF pulse. Therefore, water has a low (hypo-intense) signal and appears dark. In the brain, white matter appears white, gray matter appears gray and cerebro-spinal fluid (CSF) appears dark.

To improve contrast of certain tissue types, T_1-weighted imaging is often combined with a contrast agent, usually a gadolinium compound [133]. Gadolinium enhanced tissues appear extremely bright on T_1. This provides high sensitivity for the detection of vascular tissues.

T_2-weighted imaging

Another basic pulse sequence is T_2-weighted imaging. This sequence highlights differences in the T_2 relaxation time of tissues. A T_2-weighted image relies upon the transversal relaxation of the magnetization vector, and requires a long TE and long TR. Fatty tissue is suppressed on T_2, whereas fluid appears bright. In

the brain, gray matter has intermediate signal intensity (gray), white matter is hypo-intense compared to gray matter. CSF has high signal intensity (white).

FLAIR imaging

FLAIR is a special kind of T_2-weighted imaging sequence. It is based on an inversion recovery technique, which allows to suppress the fluid signal. By carefully selecting the inversion time (TI), the signal from any particular tissue can be nulled. This type of sequence is particularly useful in the detection of subtle changes at the periphery of the hemispheres and in the peri-ventricular region close to CSF. CSF and other fluids such as blood appear dark on a FLAIR image.

cMRI in glioma patients

cMRI has become the imaging modality of choice for managing glioma patients. Due to the nature and appearance of brain tumors, one MRI sequence is often not sufficient to fully visualize the tumor and all its subregions. In current clinical practice, different MRI sequences are combined for the diagnosis and differentiation of the tumor compartments [68]. Besides the active region of the tumor, consisting of proliferating tumor cells, other tissue compartments that may occur within the pathological region are necrosis and edema. Necrosis is the formation of dead cell tissue due to a lack of nutrients and oxygen supply, often observed in the central region of a tumor. Necrosis is a hallmark of GBMs and is not commonly seen in lower grade gliomas [178]. Vasogenic edema presents itself as an accumulation of extracellular fluid in the peri-tumoral region, due to a disruption of the blood-brain barrier.

LGGs are typically hypo-intense lesions on T_1-weighted images with limited edema and a lack of enhancement after contrast agent administration (see Figure 1.5). They appear as hyperintense on T_2-weighted and FLAIR sequences. Compared to LGGs, HGGs are often radiologically more heterogeneous and present with more severe edema. FLAIR allows outlining the whole tumor (including edema) as it shows hyper-intensity throughout the entire pathological region. In HGGs, the active tumor region often shows enhancement on T_1+contrast (T1c) images (see Figure 1.6). This is the result of vascular ingrowth into the tumor, and the fact that these newly formed tumorous vessels have a disrupted blood-brain barrier leading to leakage of contrast agent. Necrosis is also differentiated on T1c images as a low signal intensity area within the enhancing tumor rim. The edema region, surrounding the tumor, appears bright on T_2-weighted images.

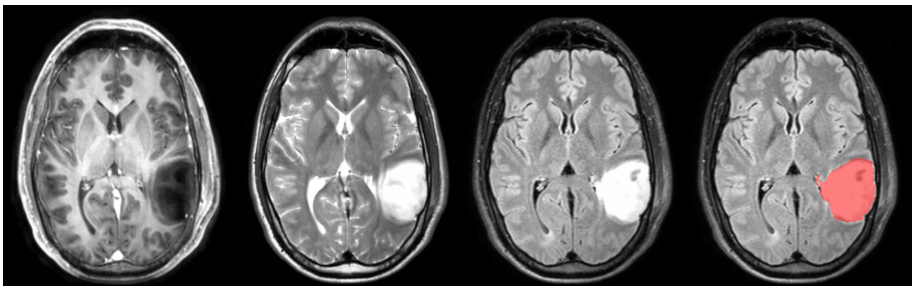

Figure 1.5: cMRI examination of a grade II astrocytoma patient. Left to right: T1+contrast, T2, FLAIR, manual delineation of the tumor.

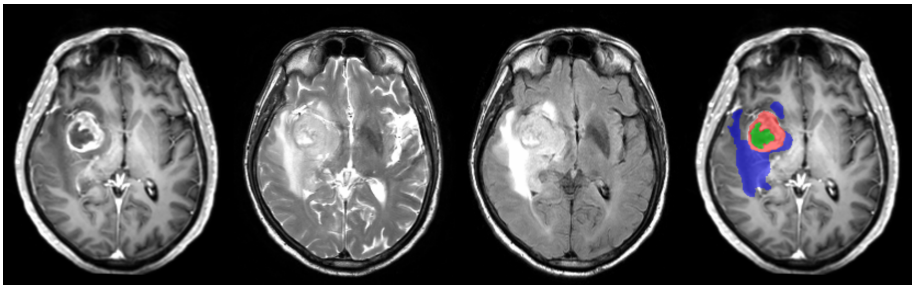

Figure 1.6: cMRI examination of a grade IV GBM patient. Left to right: T1+contrast, T2, FLAIR, manual delineation of active tumor (red), necrosis (green) and edema (blue).

Although cMRI has become an essential tool for managing glioma patients, it also comes with certain limitations. First of all, it is known that gliomas extend beyond the imageable component of the lesion, due to their infiltrative nature. cMRI tends to underestimate the full tumor extent, as it is not sensitive to infiltrated regions of peritumoral edema or healthy brain [39]. This is a major issue when considering local therapy based on target definition, such as radiation therapy. A margin around the visible tumor region is included into the target volume to take the infiltrative spread into account. As a rule, this margin is applied uniformly in all directions [203]. However, gliomas don't tend to spread in a uniform way. They preferentially invade surrounding tissue along white matter tracts, where glioma cells can be identified at considerable distances away from the gross tumor volume [58].

cMRI is not always conclusive regarding differential diagnosis of gliomas. Sensitivity of cMRI in grading gliomas has been reported to range from 55 to

83% [140]. Although most HGGs exhibit tumor enhancement, it not a reliable indicator of malignancy. Disruption of the blood-brain barrier does not allow for a reliable separation between LGGs and HGGs [85] and about one third of non-enhancing gliomas are malignant [229]. Other radiological features such as edema, mass-effect and cystic changes are also said to be predictors of higher tumor grade. However, LGGs often demonstrate one or more of these features as well [229].

Another limitation of cMRI is that it does not adequately reflect heterogeneous tissue characteristics within gliomas [109]. HGGs can exhibit different stages of the disease throughout the same lesion, with mixed biological profiles that have the capacity to develop early resistance to treatment [63]. The conventional approach of homogeneously delivering therapy over the entire tumor volume clearly seems inadequate. For instance, by delivering radiation dose homogeneously across the target volume, spatial heterogeneity of the disease is not at all considered. Alternative imaging techniques should be considered, such that increased radiation doses can be delivered to the most malignant parts of the tumor and the risk of radiation necrosis can be reduced. Advanced imaging techniques such as perfusion-weighted imaging (PWI), diffusion-weighted imaging (DWI) and magnetic resonance spectroscopic imaging (MRSI) have been suggested to guide this kind of customized radiation therapy, as they provide good association with different aspects of histology [63].

As intratumoral heterogeneity is not well envisioned by cMRI, it is an improper tool for biopsy targeting. Underdiagnosing an HGG would have the implication of depriving patients from the radical treatment options they need. Advanced MRI modalities such as PWI, DWI and MRSI have been suggested to detect the most malignant part of a tumor for biopsy sampling [22].

During follow-up, response to treatment is most commonly evaluated based on the response assessment in neuro-oncology (RANO) criteria [272]. These criteria qualify the outcome of treatment based on two-dimensional measurement of the contrast-enhancing regions together with the assessment of extent of T_2 hyperintensities. However, contrast enhancement is non-specific and often contains both residual or recurrent tumor as well as therapy-related changes. Differentiating between both phenomena has proven difficult based on cMRI. Twenty to 30% of patients treated with combined radiotherapy and chemotherapy show increased contrast enhancement within the first 3 months after completing treatment, which eventually disappears or stabilizes [25, 23]. This imaging pattern is called pseudo-progression. It is caused by a local inflammatory reaction, accompanied by edema and increased abnormal vessel permeability induced by irradiation. Pseudo-progression does not represent actual tumor progression. In fact, it is often indicative of longer survival, presumably because it represent a robust response to treatment [31].

Radiation necrosis is a potential long-term complication in patients treated with radiotherapy [24]. As opposed to pseudo-progression, radiation necrosis may appear months up to several years after radiotherapy and it implicates irreversible and progressive tissue injury. It usually occurs at the original tumor site and presents most often as contrast-enhancing lesions with central necrosis and peri-lesional edema. As such, it is often indistinguishable from tumor relapse on cMRI. The advent of anti-angiogenic therapies has given rise to another treatment-related effect called pseudo-response [53]. Pseudo-response involves a rapid decrease of contrast enhancement and edema, usually within 24 hours after the administration of the anti-angiogenic agent (e.g. bevacizumab). These early changes don't reflect a real anti-tumoral effect, but they correspond to a normalization effect on abnormally permeable tumor vessels. Simultaneously, no significant changes in the lesion pattern are found on T_2 or FLAIR images. Advanced MRI sequences such as PWI, DWI and MRSI have been shown to significantly aid in distinguishing these confounding treatment effects from true tumor response or progression, as will be discussed in the next sections.

1.2.3 Perfusion-weighted imaging

In the strict sense, cerebral perfusion is defined as the steady-state delivery of nutrients and oxygen to brain tissue, expressed as a rate per unit volume or mass. In the context of PWI, however, the term perfusion broadly refers to several tissue hemodynamic parameters that can be estimated from dynamic imaging data. It is well known that the vasculature of tumors differs from that of normal brain tissue. PWI proposes to measure the degree of tumor angiogenesis (i.e. the formation of new blood vessels from existing vasculature) and capillary permeability, both of which are important markers of tumor malignancy and prognosis. Several PWI techniques have been suggested, all of which rely on dynamic scanning upon the passage of a tracer through the cerebro-vascular system. Techniques such as dynamic susceptibility-weighted MRI (DSC-MRI) [233] and dynamic contrast-enhanced MRI (DCE-MRI) [248] exploit signal changes that are induced during passage of a contrast agent, altering relaxation times of the local tissue. Arterial spin labelling (ASL) techniques, on the other hand, magnetically label blood spins upstream from the imaging section with inverting or saturating RF-pulses [61]. This section will mainly discuss the principles and application of DSC-MRI, as it is clinically the most relevant technique and it is also used throughout this thesis.

Principles of DSC-MRI

DSC-MRI was first described by Villringer *et al.* in 1988 [261]. The technique is based on the detection of the first passage of a bolus of a non-diffusable paramagnetic contrast agent through a given tissue. When such a contrast agent (typically gadolinium-based) is injected into a vein, its paramagnetic properties locally alter the T_2^* (or T_2) and T_1 relaxation times of water. These changes in relaxation properties can be exploited to quantify haemodynamic parameters using either T_2^*- (or T_2-)weighted sequences (detected as signal loss) or T_1-weighted sequences (detected as signal enhancement). DSC-MRI is most commonly performed using a T_2^*-weighted gradient echo sequence [233]. It can also be performed by spin-echo T_2-weighted sequences, but these are not frequently used in daily clinical practice. The reduction in T_2^* is due to the signal dephasing associated with susceptibility-induced field gradients surrounding the paramagnetic contrast agent. This effect is more pronounced in areas where the contrast agent is compartmentalized since it increases the induced gradients [261]. This makes T_2^*-based techniques particularly useful for studying brain perfusion, as the contrast agent is largely confined within the blood vessels due to the presence of the blood-brain barrier.

DSC-MRI dynamically measures the signal loss throughout the brain, thereby acquiring a signal intensity-versus-time curve in each voxel. Since the transit time of the bolus through the brain is in the order of a few seconds, a rapid MRI technique is required to characterize the signal loss associated with the passage of contrast. Echo-planar imaging EPI is currently the most commonly used sequence, although this technique suffers from image distortions in regions of high magnetic susceptibility gradients (e.g. at tissue-air interfaces) [34]. Quantification of bolus tracking using an intravascular agent is based on the principles of tracer kinetics for non-diffusable tracers [291]. As MRI does not measure the tracer concentration directly, the concentration has to be inferred from the effect on the relaxation time. It has been shown empirically that, for the doses commonly used in clinical studies, the change in the relaxation rate R_2^* (which is the inverse of T_2^*) is linearly proportional to the concentration of the contrast agent [261, 214]:

$$C(t) = k\Delta R_2^*(t) = -\frac{k}{TE} ln\left(\frac{S(t)}{S_0}\right) \tag{1.1}$$

where *C(t)* represents the concentration at time *t*, *S(t)* is the signal intensity at time t and S_0 is the pre-contrast signal intensity. k is a proportionality constant that depends on the contrast agent, the tissue type, the field strength and the pulse sequence parameters [21]. Cerebral blood volume (CBV) is the most

robust and most commonly used parameter derived from DSC-MRI [39]. Zierler *et al.* have shown that, in the absence of contrast leakage and recirculation, CBV is proportional to the area under the tracer concentration-versus-time curve [291]:

$$CBV = \alpha^{-1} \frac{\int C(t)dt}{\int C_a(t)dt} \tag{1.2}$$

where α is a proportionality constant that accounts for hematocritic differences between small and large vessels [36]. $C_a(t)$ represents the tracer concentration in a voxel occupied by a large vessel, which is referred to as the arterial input function (AIF). Most often, the AIF is estimated from a major artery (e.g. the middle cerebral artery) and this global AIF is considered throughout the whole brain. However, such a global AIF neglects the spatial separation between the AIF measurement site and the voxels of interest, introducing quantification errors due to bolus dispersion [233]. In general, relative CBV (rCBV) values are often considered, by directly integrating the voxel-wise concentration-time-curves and ignoring the AIF.

In practice, assumptions of negligible recirculation and contrast agent leakage are violated. The *C(t)*-curve has a contribution of recirculation of the contrast agent, corresponding to a second, smaller and more spread peak in the curve. In brain tumors, disruption of the blood-brain barrier causes leakage of the contrast agent. This residual contrast agent locally alters T_1 and T_2^* relaxation times after passage of the tracer through the vessels, such that the baseline intensity signal does not return to its original value. Usually, signal enhancement due to T_1 shortening will mask signal loss due to T_2^* shortening [117]. To cope with this T_1 effect, it has been suggested to use a small pre-loading dose of contrast agent, such that T_1 shortening is initially induced and its effect is negligible when the second bolus is used to quantify perfusion [39]. However, the use of such a pre-loading dose does not overcome the decreased T_2^* effect due to the loss of tracer compartmentalisation.

Alternative post-processing strategies have been developed to compensate for recirculation and leakage effects. A common approach is to fit a gamma-variate function, which approximates the curve that would have been obtained without recirculation or leakage [269]. This non-linear fitting procedure is typically unstable in portions of the image with low signal-to-noise ratio [158]. More robust approaches to differentiate true rCBV involve generating a model function of the signal as the sum of 2 opposing T_1 and T_2^* effects and then fitting a linear model to estimate the properties of each component [270]. A diagram illustrating the calculation of rCBV and all involved steps is shown in Figure 1.7.

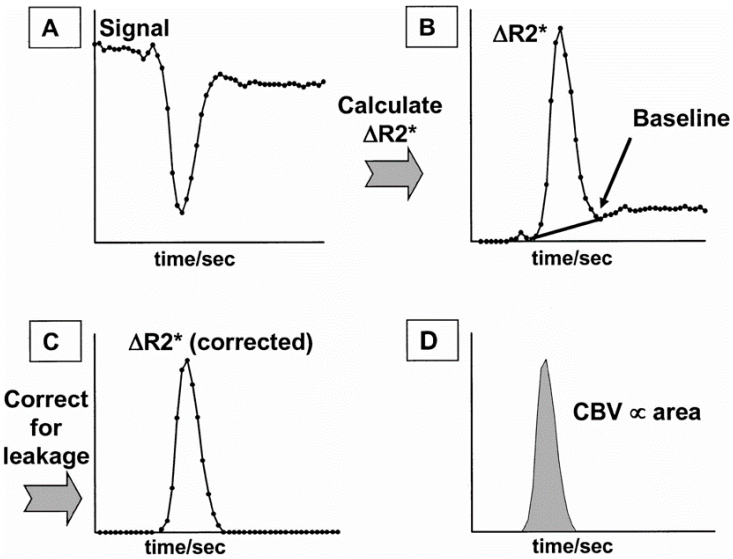

Figure 1.7: A: Signal intensity decrease during passage of contrast agent bolus is measured using DSC-MRI. B: Change in the relaxation rate (ΔR_2^*) is calculated from signal intensity, and a baseline subtraction method is applied to measured data. C: Corrected ΔR_2^* curve. D: rCBV is proportional to the area under curve. *Source: Cha et al., Radiology, 2002 [40].*

According to indicator dilution theory, the concentration of tracer in the tissue can be described by [36]:

$$C(t) = \alpha \cdot CBF \int_0^t C_a(\tau) R(t-\tau) d\tau \tag{1.3}$$

where α is the same proportionality constant as in Equation (1.2), CBF represents cerebral blood flow and *R(t)* is the residue function, which describes the fraction of contrast agent remaining in the tissue at time t. The integral operator in Equation (1.3) represents a convolution, accounting for the non-instantaneous arrival of the bolus. In order to calculate CBF, Equation (1.3) must be deconvolved in order to isolate the residue function scaled by CBF [35]. Mathematical deconvolution is sensitive to noise, and different methods have been proposed [183]. The most commonly used method is to rewrite Equation (1.3) as a matrix equation and invert the resulting system using singular value decomposition. The mean transit time MTT is defined as the average time for a molecule of contrast agent to pass through the tissue vasculature following an

ideal instantaneous bolus injection. MTT can be derived from CBV and CBF through the central volume theorem [243]:

$$MTT = \frac{CBV}{CBF} \tag{1.4}$$

Several other parameters can directly be obtained from the concentration-time curve: bolus arrival time (BAT), time to peak (TTP), maximum peak concentration (MPC), area under the peak (AUP) and full width half maximum (FWHM) [233]. Although the analysis of these parameters is straightforward and does not require estimating the AIF nor assuming an intact blood-brain barrier, these parameters do not provide direct measures of perfusion, and they are strongly influenced by inter-individual differences in micro-vascular structure, cardiac output, injection technique and contrast agent leakage [191]. Therefore, these parameters are not often used in clinical practice.

PWI in glioma patients

PWI is increasingly being used as a diagnostic tool assessing regional variations in cerebral micro-vasculature in normal and diseased brain [22, 39]. With relatively short imaging and processing times, PWI is a promising tool that could be incorporated as part of the routine clinical evaluation of brain tumors. CBV measurements can be used as an adjunct to cMRI to help assess the degree of neo-vascularization, evaluate tumor grading and malignancy and identify tumor-mimicking lesions such as cerebral abscesses or radiation necrosis.

Like other tumors, malignant gliomas have the remarkable ability to promote vascular ingrowth, thereby recruiting vascular networks for their further growth and proliferation [40]. The degree of neo-angiogenesis is one of the main histologic features for determining glioma grade of malignancy [157]. In astrocytomas, the strong correlation between rCBV and tumor grade has been extensively documented [3, 29, 122, 246], with maximal rCBV values increasing along with tumor grade. One important caveat concerns pilocytic astrocytomas, which may also present with high blood volumes due to vessel dilation. Oligodendrogliomas, on the other hand, have delicate neo-angiogenic vessels, and they demonstrate elevated rCBV levels, even in LGGs, that can be as high as in GBMs [41]. HGGs also tend to show higher capillary leakage, although the correlation with grading was found to be weaker than with rCBV measurements [139]. It is important to recognize that there is an overlap of rCBV values between different glioma grades. Moreover, due to the inherent histologic heterogeneity of gliomas, CBV measurements can vary considerably. HGGs may focally show an extreme angiogenic response, whereas vascular

density is in the range of that of normal white matter in other regions [273]. Therefore, rCBV maps should be interpreted in combination with cMRI, which can provide other valuable information such as integrity of the blood-brain barrier and characteristics of T_2 abnormality.

Metastatic brain tumors often do not pose a diagnostic dilemma on cMRI, because they tend to be multiple, they are located near the gray matter-white matter junction and they often present with a known history of systemic malignancy. On the other hand, approximately 30% or more of metastatic brain tumors manifest as a single lesion, and in some cases as the initial clinical presentation of malignancy [197]. cMRI characteristics of solitary metastasis and GBM are non-specific and cannot always differentiate between the two. Both tumor types are highly vascular with rCBV values, and mixed results have been reported in the ability of differentiating the two based on intra-tumoral rCBV values [22, 40]. However, DSC-MRI has been shown useful in differentiating peri-tumoral edema of primary gliomas and metastatic tumors. rCBV values within the peri-tumoral edema of metastases were significantly lower than those within the infiltrative edema of GBMs [137]. This is explained by the fact that in HGGs, the peri-tumoral region consists of a variable combination of vasogenic edema and infiltrating tumor cells, whereas metastatic tumor are not infiltrative.

Biopsy remains the definitive method to determine tumor type and grade. It is well known that sampling errors occur in biopsies of HGGs, partly due to the high intra-tumoral heterogeneity [276] which is not well reflected in cMRI. DSC-MRI has been suggested to guide stereotactic biopsy [40, 268]. rCBV values have been shown to correlate strongly with tumor cell density and microvascular density in gliomas, based on histology [3, 217]. Maximal rCBV values have also been correlated with vascular endothelial growth factor [162]. As such, rCBV maps may be helpful in biopsy targeting, to evaluate the most malignant area of the tumor and thereby avoiding underdiagnosis.

The predictive value of PWI for the prognosis of glioma patients has been well documented. rCBV values have been reported to be a better predictor of survival than the degree of contrast enhancement [150]. Law *et al.* suggested that rCBV values in LGGs better correlated with time to progression than initial histopathological grading of the tumor [138]. During treatment, early changes in PWI parameters have been able to indicate response to treatment [37] and to differentiate between tumor progression and regression [107]. PWI has also been shown helpful in the differentiation between tumor progression and radiation necrosis. CBV maps can demonstrate the pathologic differences in vascularization, as radiation necrosis demonstrates reduced CBV values due to vascular dilation and endothelial injury [8], whereas elevated CBV is found in progressing tumor due to neo-angiogenesis.

1.2.4 Diffusion-weighted imaging

DWI probes the microscopic motion of water molecules within tissue. This diffusion process is driven by thermal agitation, following the Brownian law [144]. Molecular diffusion in tissues is not free, but reflects interactions with many obstacles, such as macromolecules, fibres, and membranes. As such, DWI can provide insight into the structural and organizational features of healthy and impaired biological tissues. DWI is currently the only method for the in-vivo and non-invasive quantification of water diffusion. Over the last decades, new diffusion models have been proposed to obtain quantitative diffusion measures. Diffusion tensor imaging (DTI) is a fairly recent diffusion MRI method, modeling the three-dimensional nature of the diffusion process [12]. Diffusion kurtosis imaging (DKI) is a very recent technique, developed as an extension from DTI, providing information on diffusional heterogeneity and micro-structural complexity [113].

Principles of diffusion-weighted MRI

Le Bihan *et al.* were the first to publish on diffusion-weighted images of the human brain in 1986 [145]. Diffusion-weighted images are raw images where the signal has been sensitized to diffusion by strong magnetic field gradient pulses [143]. In the presence of a spatially varying magnetic field, random motion of water molecules results in a dephasing of the MR signal, as the precession speed of individual spins will vary with the local variation in magnetic field strength. By applying these large magnetic field gradients in particular directions, diffusion can be made the dominant image contrast mechanism. The most common approach to measure diffusion is to use a spin-echo acquisition scheme [242]. Strong rectangular gradient pulses are added to the sequence, equal in duration and magnitude but opposite in phase (see Figure 1.8). In the absence of diffusion motion, the dephasing caused by the first pulse is fully reversed by the second pulse, such that the phase of stationary spins is unaffected. However, if the water molecules have moved between the 2 pulses, the rephasing of the MR signal is incomplete and the measured signal is decreased.

Assuming that the molecular diffusion is purely random, the measured MR signal can be modeled by the following equation [242]:

$$\frac{S(TE)}{S_0} = \exp(-\gamma^2 G^2 \delta^2 (\Delta - \frac{\delta}{3}) D) \tag{1.5}$$

where *S(TE)* represents the acquired MR signal, S_0 is the signal without diffusion weighting, γ is the gyro-magnetic ratio, G is the strength of the gradient pulse,

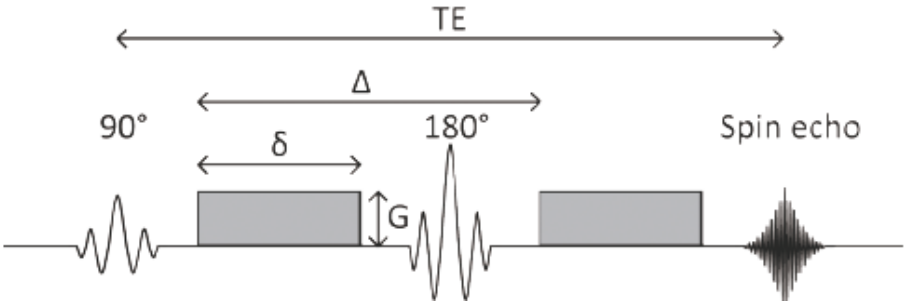

Figure 1.8: Diffusion-sensitized spin echo sequence. The spin echo sequence is modified by the addition of two gradient pulses (grey blocks) which induce and then reverse a spatially-dependent phase shift. *Source: Winston et al., Quant Imaging Med Surg, 2012 [278].*

δ is the duration of the pulse, Δ is the time between the 2 pulses and D is the diffusion coefficient. Le Bihan *et al.* reformulated this expression, combining all diffusion-sensitizing parameters in the so-called b-value [145]:

$$\frac{S(TE)}{S_0} = \exp(-b \cdot ADC) \tag{1.6}$$

Compared to Equation (1.5), the diffusion coefficient D has been replaced by the apparent diffusion coefficient ADC, to indicate that the diffusion process is not free in tissues, but hindered and modulated by many mechanisms (restriction in closed spaces, tortuosity around obstacles, etc.) and that other sources such as blood flow in small vessels or cerebrospinal fluid in ventricles also contribute to the signal attenuation. In diffusion-weighted images, signal attenuation gets stronger with higher diffusion and higher b-values. However, these DWI images are still sensitive to T_1 and T_2 relaxation contrast, which can confound interpretation of the images (e.g. the well-known T_2 shine-through effect). Hence, if possible, truly quantitative diffusion images are preferred. Based on Equation (1.6), the ADC can be calculated by collecting images with at least 2 different b-values:

$$ADC = \frac{ln\left(\frac{S_1}{S_2}\right)}{b_2 - b_1} \tag{1.7}$$

In a homogeneous fluid, diffusion is similar in all spatial directions (termed "isotropic diffusion"). However, in biological tissue, extent of diffusion is also affected by the presence of cellular structures which act as barriers to diffusion. Water molecules will thus be impeded in their movement by natural barriers, such as macro-molecules or cell membranes. In tissue with a highly organized structure, diffusion may be hindered more strongly in certain directions than in others (termed "anisotropic diffusion", see Figure 1.9). White matter, for instance, is characterized by a high degree of anisotropy, with high diffusion along the direction of the axonal fibres, and low diffusion perpendicular to it. Obviously, it might not be sufficient in certain tissue structures to represent the diffusion process in only one direction. Diffusion tensor imaging (DTI) allows to model this orientational dependency of water diffusion throughout the brain [12]. In the DTI model, anisotropic diffusion is characterized by a diffusion tensor. This tensor can be calculated based on at least 6 independent measurements with diffusion gradients along six non-collinear directions, using b-value 0 and a b-value in the range of 800-1000sec/mm^2 for clinical applications. DTI models 3-D diffusion by an ellipsoid, with the principal axes of the ellipsoid derived from an eigenvalue decomposition of the diffusion tensor. Based on the eigenvalues of the ellipsoid, different DTI scalar measures can be quantified. Fractional anisotropy (FA) quantifies the degree of diffusion anisotropy, scaling from 0 (isotropic diffusion) to 1 (diffusion in only one direction). Mean diffusivity (MD) represents the magnitude of diffusion averaged over the applied gradient directions. Axial diffusivity (AD) and radial diffusivity (RD) reflect the magnitude of diffusion along the principal axis of the ellipsoid and in a plane perpendicular to the principal axis, respectively.

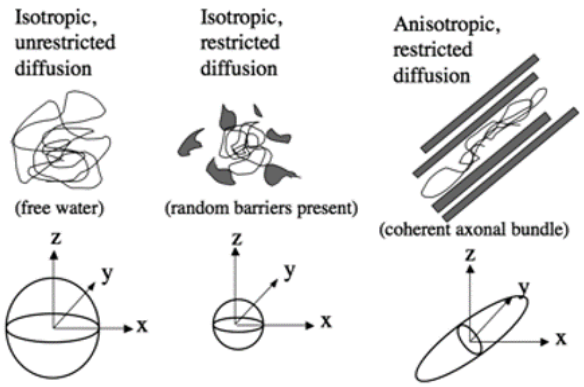

Figure 1.9: Isotropic unrestricted and restricted diffusion and anisotropic diffusion. *Source: Mukherjee et al., Am Soc Neuroradiology, 2008 [173].*

Both DWI and DTI assume a mono-exponential relation between the signal intensity and the b-value, as defined in Equation (1.6). This relation holds as long as water diffusion is a random process, which can be modeled by a Gaussian distribution [5]. At higher b-values, corresponding to longer diffusion times, an increasing fraction of molecules encounter obstacles which restrict their diffusion. Non-Gaussian diffusion is readily observed at b-values larger than $1500 sec/mm^2$ [12]. Diffusion kurtosis imaging (DKI) approximates the diffusion-weighted signal more accurately by taking into account the degree of non-Gaussian diffusion [196]. DKI extends the exponent of the DWI/DTI model with a second-order term. The signal attenuation can more accurately be modeled by the following equation:

$$ln\left(\frac{S(TE)}{S_0}\right) = -b \cdot D + \frac{1}{6}b^2D^2K \tag{1.8}$$

where the kurtosis K is a dimensionless metric which quantifies the deviation from Gaussian behaviour. The DKI model is fully defined by 22 elements: a non-diffusion-weighted signal b_0, 6 parameters to characterise the diffusion tensor and 15 parameters to characterise the kurtosis tensor. To fit the DKI model, at least 15 non-collinear gradients need to be acquired with 2 non-zero b-values in the high range [196]. Previous studies have shown that maximum b-values in the range of 2000 to $3000 sec/mm^2$ are required to obtain sufficient sensitivity to the non-Gaussian diffusion component [113]. In analogy with MD in DTI, the mean kurtosis (MK) is defined as the averaged kurtosis over all measured diffusion directions. Axial kurtosis (AK) represents the kurtosis coefficient in the direction of main diffusion, and radial kurtosis is the mean kurtosis value from all directions perpendicular to the direction of main diffusion. DKI also provides the conventional diffusion metrics, and several authors have claimed that DKI quantifies these metrics more accurately than the DTI model [105, 257].

Diffusion-weighted MRI in gliomas

Diffusion metrics are non-specific, and it is often difficult with the current modeling techniques to relate changes in diffusion measures directly to specific physiological changes at the molecular level [143]. Many variables such as viscosity and cellular shape, size, density and permeability can influence diffusion. Nevertheless, ADC levels in gliomas have been commonly related to cellular density [39, 206]. In fact, a significant inverse correlation between minimum ADC values and cellular density based on histology has been shown [92, 217]. Several studies reported that glioma grade correlates inversely with minimum ADC, most

likely because of increasing tumor cellularity with grade [28, 121, 125]. However, a substantial overlap in ADC values between grades does exist, precluding its use for differentiation in individual patients. MK values have been shown to better separate between glioma grades than ADC (or MD) and FA [206, 253]. MK values increase with tumor grade, probably reflecting a higher degree of tissue complexity.

Price *et al.* reported better delineation of glioma margins by using DTI compared to cMRI, based on correlation with histologic examination of image-guided biopsies [203]. Kallenberg *et al.* found significant differences in ADC and FA values in corpus callosum which was either unaffected or affected with infiltrating tumor cells [116]. Another study used MD and FA values from both the tumor region itself as well as from the peri-tumoral edema to successfully differentiate GBMs from metastases [32]. Other studies, however, reported that ADC values did not allow a clear demarcation between bulk tumor and peri-tumoral edema [70, 189].

DTI has had a major impact on intra-operative guidance in tumor resection, based on FA maps and white matter tractography [284]. Different patterns of FA maps have been described to differentiate white matter tracts that are displaced or distorted but intact from those that are infiltrated or destroyed by neighboring tumor [111]. DTI tractography maps can be overlaid onto high-resolution cMRI images to identify the location of eloquent white matter tracts displaced by tumor and to predict the degree of post-operative functional impairment based on surgical injury to these tracts [86, 142].

After surgical resection, acute cellular damage commonly causes ischemia and infection at the margin of the resection bed. This infarcted region will often demonstrate contrast enhancement at the early post-operative stage, mimicking tumor progression [239]. Immediate post-operative DWI will show focally reduced diffusion in this area, related to acute intra-cellular swelling [39]. Detection of this early post-operative diffusion abnormality is important, as it can help in differentiating such pseudo-progression from actual tumor recurrence. DWI has also shown its added value in monitoring treatment response. Successful radiation- and/or chemotherapy will result in reduced tumor cellularity and induce necrosis, characterized by increased ADC values [43, 100].

1.2.5 Magnetic resonance spectroscopic imaging

Whereas cMRI essentially provides anatomical information regarding tissue, magnetic resonance spectroscopy (MRS) assesses the concentration of certain biochemical components, called metabolites, in the tissue under investigation. MRS is sensitive to the differences in resonance frequency of the different

metabolites due to the so-called chemical shift effect. As such, MRS provides spectra exhibiting peaks which are proportional to the metabolites' concentrations. Whereas MRS provides biochemical information on a specific location (also called single-voxel MRS), magnetic resonance spectroscopic imaging (MRSI) reveals the spatial distribution of metabolites throughout a region of interest.

Principles of MRS

The basic principle that enables MRS(I) is the so-called shielding effect or chemical shift effect. When an atom is placed in a magnetic field, its electrons will circulate around the direction of the applied magnetic field, thereby creating a small opposing magnetic field. One could say that the electron cloud "shields" the atom from the external magnetic field B_0, thereby locally reducing the effective magnetic field B_{eff}:

$$B_{\text{eff}} = B_0(1 - \sigma) \tag{1.9}$$

σ is called the shielding constant, a dimensionless parameter that depends on the electrical environment of the nucleus. The precession rate of the magnetic spin is proportional to the effective magnetic field:

$$f_{\text{eff}} = \frac{\gamma(1 - \sigma)B_0}{2\pi} \tag{1.10}$$

The degree of shielding (i.e. the magnitude of σ) depends on the chemical compound in which the nucleus is located, as well as the specific location of the nucleus within the compound. As a result, the observed resonance frequency f_{eff} will be slightly different for different metabolites. This so-called "chemical shift" from the reference frequency f_0 allows to differentiate between the metabolites. The chemical shift can be expressed in Hz or in parts-per-million (ppm). The latter is commonly preferred because it is independent from the spectrometer frequency. Figure 1.10 shows an exemplary spectrum acquired within a GBM. The different resonant peaks correspond to different metabolites. The area under a peak is approximately proportional to the metabolite's concentration.

As is the case for cMRI, MRS equipment can be tuned to pick up signals from different nuclei within the body. Proton ^{1}H-MRS is most commonly used, assessing the chemical shift of hydrogen atoms. Many pulse sequences used for cMRI can be used for spectroscopy as well. However, in MRS water suppression is required as the relative concentration of water to metabolite

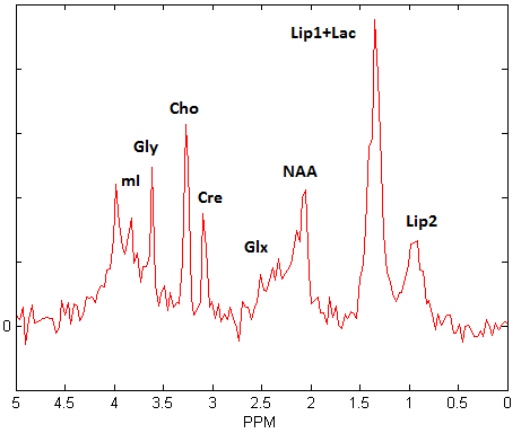

Figure 1.10: Short-TE ^{1}H-MRS spectrum acquired within a GBM. The main metabolites of interest are indicated, right to left: lipids at 0.9ppm (Lip2), lactate (Lac), lipids at 1.3ppm (Lip1), N-acetyl-aspartate (NAA), glutamate+glutamine (Glx), creatine (Cre), choline (Cho), glycine (Gly) and myo-inositol (mI))

is roughly 10,000 to 1. Several methods exist based on the frequency and T_1 relaxation of water, such as chemical shift selective suppression (CHESS) [93] and multiply optimized insensitive suppression train (MOIST) [180]. An important choice in the MRS protocol is the duration of the echo time TE. Long TEs ($\geq$ 130ms) allow an easier extraction of metabolic information of the slowly decaying components, like choline (Cho), creatine (Cre) and N-acetyl-aspartate (NAA). Short TEs ($\leq$ 50ms) result in larger peaks in the MRS spectrum and provide more metabolite information as there has been less signal decay [277]. This allows for instance the observation of glutamine (Gln), glutamate (Glu) and myo-inositol (mI). In short TE signals macromolecules are also prominent, which are very broad peaks that contribute to an underlying and partially unknown baseline. Single-voxel MRS measures a signal from a volume limited to a single voxel. It is preferred in case of localised and/or homogeneous pathology, as it provides better quality data at shorter acquisition times. MRSI, also called chemical shift imaging (CSI), is able to acquire multiple signals from a (2D- or 3D-)grid of voxels. While MRSI requires longer acquisition and processing times, it has the advantage of providing spatial distribution of the metabolites (see Figure 1.11). Due to the inherently low signal-to-noise ratio in MRS(I), spatial resolution is much lower than in other MRI modalities. For instance, a voxel size of $1 \times 1 \times 1 cm^3$ is common in MRSI. For the same reason, MRSI is restricted to a limited region of interest, as acquisition times would become too long when measuring full brain slices. However, recent advances in hardware

and pulse sequences have made it possible to scan much larger brain volumes with good signal-to-noise ratio and in reasonable scan time [185].

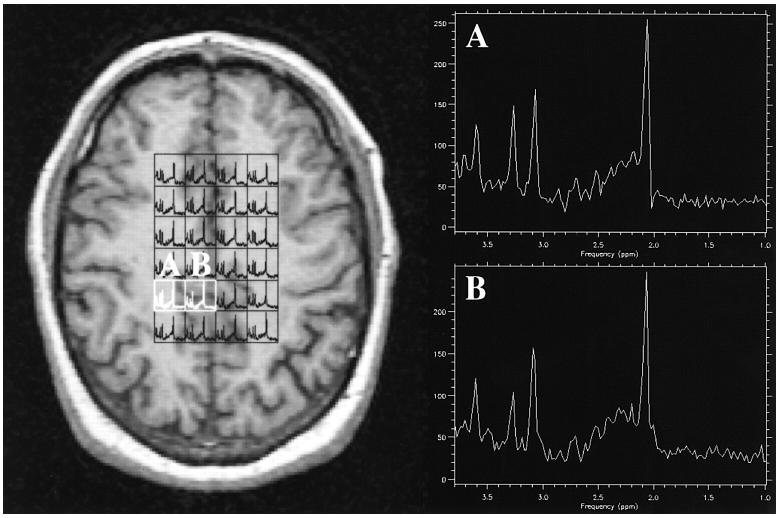

Figure 1.11: T_1-weighted structural image overlaid with the MRSI grid. Detailed spectra are shown for a normal-appearing white matter voxel (A) and for a cortical gray matter voxel (B). *Source: Chard et al., Brain, 2002 [44].*

The concentration of a metabolite is linearly proportional to its spectral peak area. However, peak area measurements are complicated by resonance overlap, baseline distortions and non-ideal lineshapes. Metabolite quantification is the process of extracting (relative) metabolite levels from the acquired MRS(I) signals. Figure 1.12 gives an example of metabolite maps obtained through quantification from a short-TE MRSI acquisition. Various methods have been used to measure peak areas, ranging from simple integration to fitting algorithms in the time- or frequency domains [199, 204, 207]. A common approach is to model the MRS signal based on a selected set of metabolite profiles from a database. This allows to implicitly incorporate prior knowledge that relates to the individual peaks in the spectrum. The metabolite levels are estimated from the weighting coefficients (i.e. amplitudes) of the linearly combined metabolite profiles from the database. The quantification model should allow for small profile corrections in spectral parameters as frequency shifts, damping corrections and phase shifts, since these parameters might undergo small variations across different measurements [208]. The presence of a residual water signal and a macromolecular baseline (in short-TE MRS) complicate the quantification procedure. These confounding factors have to be taken into account by the

quantification method, either in a pre-processing step [199] or as part of the quantification procedure [207].

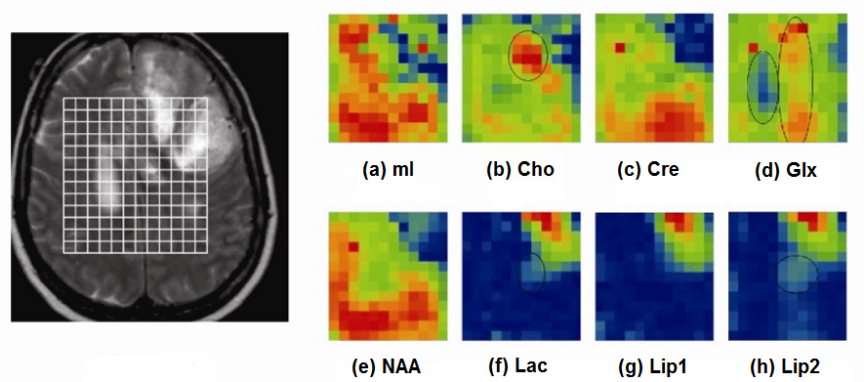

Figure 1.12: Left: T_2-weighted image of a GBM patient overlaid with the MRSI grid. Right: quantified metabolite maps obtained from short-TE ^{1}H-MRSI acquisition. (a)myo-inositol, (b)choline, (c)creatine, (d)glutamine+glutamate, (e)N-acetyl aspartate, (f)lactate, (g)lipids at 1.3ppm, (h)lipids at 0.9ppm.

Metabolites in the human brain

This section discusses the main metabolites that can be detected by MRS(I) at clinical field strengths on short- and long-TE spectra in the healthy brain as well as in gliomas. These metabolites will be the subject of research throughout this thesis. For an exhaustive list of all compounds that have been detected in the human brain, the reader is referred to [89].

The main metabolites of interest are shown on the short-TE spectrum in Figure 1.10. They are discussed below in order of appearance on the ppm axis. Lipid resonances occur at 0.9ppm (Lip2), at 1.3ppm (Lip1) and at 2.8ppm. Lipids are reliable biomarkers of tumor progression, as they indicate increased concentrations of necrosis [195], which is a histological hallmark of GBM. Concentrations of mobile lipids have been reported to correlate linearly with the amount of necrosis [45], and therefore seems to reflect the severity of tissue damage.

Lactate (Lac) presents as a doublet (2 resonant peaks) at 1.33ppm, which closely overlaps with Lip1, and a quartet at 4.09ppm. Lac is the end product of anaerobic glycolysis, and is usually detected when the energy metabolism is severely affected [78]. Malignant transformation of gliomas involves an increase

in cellular density which is accompanied by ischemia, resulting in higher Lac levels [30].

N-acetyl aspartate (NAA), with its main resonance peak at 2.01ppm, is the largest signal in the normal adult brain spectrum. As it is regarded as a marker of neuronal density, reduced NAA levels may reflect neuronal loss or disfunction [167]. NAA levels have been found to significantly decrease with increasing glioma grade [241].

Glutamate (Glu) and glutamine (Gln) are strongly overlapping, both having multiplets in the regions of 2.05ppm, 2.40ppm and 3.75ppm. They are often reported as a composite (Glx). Glu and Gln are important amino acids in the brain and they are involved in neuro-transmitter synthesis [240].

Creatine (Cre) occurs in the MRS spectrum as 2 singlets, at 3.03ppm and at 3.9ppm. As the resonance peak at 3.03ppm synthesizes both Cre as well as phosphocreatine, quantification of Cre is also referred to as "total creatine". Cre is a marker of energy metabolism [279]. Cre is reduced or even absent in brain tumors. This decrease in Cre level has been related to the higher metabolic demands of tumorous tissue [30], although its significance has not been fully explained so far.

Choline (Cho) is a composite peak at 3.2ppm, consisting of contributions from the tri-methylamine groups of glycerophosphocholine (GPC), phosphocholine (PC), and a small amount of free choline itself [10]. It is therefore also referred to as "total choline". These compounds are involved in membrane synthesis and degradation, and they are often elevated in disease states where increased membrane turnover is involved, such as tumors [215].

Glycine (Gly) has a resonant peak at 3.55ppm. It is usually not detected in healthy brain due to its low amplitude [30]. Gly is an amino acid that is involved in several metabolic pathways. Elevated Gly levels have been reported in GBM and it may be a marker of tumor malignancy [155, 249].

One of the larger signals in short-TE spectra comes from myo-inositol (mI), a multiplet with its main resonances in the range of 3.5-3.7ppm. Due to the strong overlap, Gly is sometimes quantified as part of the mI complex [155]. mI is considered to be essential for cell growth, acting as an osmolite and a storage form of glucose. It is reported absent or low in non-glial tumors like meningiomas and in high levels in LGGs [89]. mI has also been reported to play an important role in glioma invasion [30].

MRS(I) in gliomas

As gliomas, and especially HGGs, are known to be highly heterogeneous, MRSI examination is usually preferred over single-voxel MRS as it allows to assess spatial variability of the tumor characteristics. Detection of significant differences in metabolite levels can increase sensitivity in determining glioma grade. LGGs are generally characterized by a relatively high concentration of NAA, a low level of Cho and absence of Lac and Lip. Progression in glioma grade is reflected in the progressive decrease in NAA and mI levels and an elevation of Cho up to grade III [30]. Malignant transformation is accompanied by the presence of Lac, indicating hypoxic tumor metabolism. Lip is a hallmark of GBM, indicating the presence of necrosis, although elevated Lip levels have also been found in HGGs with no macroscopic necrosis based on cMRI [249]. Both in LGGs as well as in anaplastic gliomas, the normalized level of Cre has been shown to be a prognostic factor of progression-free survival time, with elevated Cre levels indicating earlier progression [98]. As normalization with respect to normal appearing white matter is not always possible due to the limited region of interest, metabolite ratios are commonly considered as a means of normalization [126]. Several studies have reported significant differentiation of LGGs and HGGs, mainly based on metabolite ratios of Cho/NAA, Cho/Cre and (Lip+Lac)/Cre [6, 76, 126]. Although considerable overlap exists among glioma grades, clinical diagnosis might benefit from incorporating MRSI to improve grading accuracy.

Assessment of intra-tumoral heterogeneity has led to the consideration of MRSI for guiding stereotactic biopsy. By targeting areas of metabolically active tumor on the basis of increased Cho or Cho/NAA levels, several studies have reported improved diagnostic accuracy compared to cMRI-guided sampling [46, 95].

MRSI has been found useful in differentiating solitary metastases from GBMs by measuring peri-tumoral Cho levels [137]. Significantly elevated Cho/Cre ratios were found in the peri-tumoral edema of GBMs, whereas no elevation was found in metastases. Other studies have reported abnormal metabolic activity beyond the radiological boundaries of the lesion, suggesting that MRSI might be able to detect tumor invasion [193, 212]. It must be noted, however, that MRSI probably has limited sensitivity with regard to microscopic infiltration, due to its low spatial resolution.

One should be cautious in differentiating between tumor recurrence and radiation necrosis upon the inspection of individual metabolites. The presence of Lip or Lac peaks is not useful, as they are present in both cases. A temporary increase in Cho-containing compounds may result from cell damage and astrogliosis after radiation therapy [103]. Complete absence of NAA and Cho peaks, or

their progressive decrease on serial MRSI, in combination with elevated Lip and Lac peaks is indicative of radiation injury. A significant increase in Cho and a decrease in NAA over time, with a consequent increase of the Cho/NAA ratio is a sensitive indicator of tumor recurrence in the appropriate cMRI context [91, 194]. Several studies found that Cho/NAA and Cho/Cre ratios are significantly higher in recurrent tumor than in radiation necrosis [238, 274]. Overall, MRSI has been shown to be a useful adjunct to cMRI in the discrimination of tumor recurrence from radiation necrosis.

1.3 Brain tumor segmentation

1.3.1 Segmentation in treatment planning and follow-up

Tumor segmentation involves the imaging-based delineation of a tumor and possibly of its subcompartments (i.e. active tumor, necrosis and edema). In gliomas, segmentation plays an important role in treatment planning as well as during follow-up, to monitor response to treatment. Targeted therapies, such as radiation therapy, high-intensity focused ultrasound and proton therapy require the definition of a localized target volume. In current clinical practice, an outline of the gross tumor volume is considered, and therapy is delivered homogeneously over this target volume. With the recent emergence of the dose painting concept in radiation therapy, segmentation of the active tumor region would allow focussing the radiation dose to the most active part of the tumor [63]. This approach is expected to increase efficacy of treatment, while at the same time reducing toxicity. Criteria to assess response to treatment in gliomas have recently been updated by the response assessment in neuro-oncology (RANO) working group [272]. Following these criteria, growth or shrinkage of the lesion is evaluated by measuring the maximal diameters of the lesion in 2 orthogonal directions. This is a rather basic assessment of tumor size, and the RANO group has acknowledged that volumetric measurements might be favourable compared to cross-product diameters in the case of irregularly shaped tumors, multi-focal tumors and tumors with cystic or necrotic components. Furthermore, several studies have shown that volumetric measures of the tumor subcompartments carry prognostic information. Pre-operative extent of enhancing tumor, necrosis and edema were found to be significant prognostic factors of overall survival time [97, 130]. The combination of pre-treatment volumes of enhancing tumor and FLAIR abnormality was found to be predictive of progression-free survival time and overall survival time in GBMs [71]. These findings suggest that inclusion of volumetric measures of the tumor subcompartments might improve prediction of individual patient survival.

Segmentation of gliomas is a challenging task. Gliomas and especially HGGs are known to be highly heterogeneous. Several stages of the disease can occur within the same lesion [190] and boundaries between pathological tissue regions are unclear, as viable tumor might present with focal necrosis [156] and peritumoral edema might be invaded by tumor cells [137]. Additionally, clinical datasets are often acquired in a highly anisotropic way, such that intra-slice resolution is much higher than inter-slice resolution. This causes problems with partial-volume effects for the segmentation. As several MRI images have to be combined in order to differentiate the tumor subcompartments, lack of co-registration between the images further complicates the segmentation process. Manual segmentation is currently the gold standard, but it is a tedious and time-consuming task, and it is susceptible to subjective interpretation. Intra- and inter-observer variabilities of 20% and higher have been reported for manual tumor delineation [168, 271]. For these reasons, volumetric delineation of the tumor and its subcompartments has so far not become widespread in clinical practice. To overcome these limitations and due to the high clinical relevance, there has been an increasing interest in the machine learning community for automated brain tumor segmentation.

1.3.2 Segmentation methods

In recent years, significant advancements have been made in the field of automated brain tumor segmentation. Segmentation methods are mainly categorized into 2 groups: region- or edge-based methods, which rely on a deformable model, and classification methods, which rely on voxel-wise intensity features. Deformable models make use of regional characteristics or edge detection in the images [166]. Active contour models use an energy minimizing, deformable spline curve which is influenced by constraint and image forces that pull it towards object contours and internal forces that resist deformation. In many cases, a level-set is evolved towards the tumor boundary by searching for the largest gradient in the image or by employing region properties. The active contour requires proper initialization near the tumor boundary, which is usually based on user input [99]. Wang *et al.* proposed a fluid vector flow model to evolve an active contour towards the tumor edge on T_1-weighted images [266]. Ho *et al.* derived a tumor probability map by applying a Gaussian mixture model to the difference image of T1c and T_1, which was used to automatically initialize the active contour [102]. Rexilius *et al.* initialized a region-growing algorithm with a tumor probability map which was obtained from a multi-spectral histogram model adaptation [210]. An issue of active contour models is that they are highly dependent on image gradients, and they might fail if the tumor region does not have well-defined borders. Also, strong image gradients

from surrounding objects may attract the active contour in the wrong direction. Atlas-based deformable models map the patient's imaging dataset to a brain atlas of healthy subjects. Tumors may be modeled as outliers relative to the expected shape [286] or signal intensity [79] of the healthy tissues. Prastawa *et al.* used a normal brain atlas to obtain healthy tissue priors, and added a tumor prior based on the assumption of tumor enhancement [202]. These priors were then combined with MRI features into an expectation-maximization model to derive tissue classification. Atlas-based methods might have difficulties in recognizing deformed normal tissue regions, as a result of the mass effect induced by a growing lesion or in the presence of resection cavities. Region- or edge-based methods generally only provide a binary separation of healthy and pathological tissue, and they require additional processing steps to differentiate the tumor subcompartments.

Classification methods, which rely primarily on voxel intensity-based features for tissue segmentation, have lately been receiving most attention [168]. They have the advantage of being able to deal with multi-parametric imaging datasets, combining voxel-wise information from different MRI sequences. Classification methods can be further categorized into supervised and unsupervised methods. Supervised classification methods naturally incorporate large amounts of prior knowledge in the form of a training dataset with known tissue labels. From this training dataset, the algorithm learns decision boundaries between tissue classes in high-dimensional feature space, which can then be applied to unlabelled test data. The most popular methods include random forests [228], support vector machines [227] and neural networks [160]. Zikic *et al.* used random forest classification with context-aware features and an additional generative model as an input, to separate tumor subcompartments from multi-parametric images [292]. Bauer *et al.* combined a random forest classifier with conditional random field regularization, to promote spatial smoothness of the tissue labels [13]. Ruan *et al.* applied support vector machines to cMRI data to monitor changes in tumor volume during follow-up [216]. Supervised classification methods require extensive training datasets with a uniform acquisition protocol, to account for the wide heterogeneity in tumor appearance among glioma grades.

Unsupervised classification methods reveal structure in a dataset by modeling the similarity within the data itself. These methods can directly be applied to any imaging dataset, irrespective of the acquisition protocol and without the need for any training data. Popular approaches include fuzzy C-means clustering (FCM) [73, 192], Gaussian mixture modeling (GMM) [115, 169] and hidden Markov random fields (HMRF) [175, 288]. FCM was introduced by Phillips *et al.* [192] and was later combined with knowledge-based techniques for improved performance [73, 119]. Juan-Albarracin *et al.* have recently applied several unsupervised methods to the cMRI dataset of the Multimodal Brain

Tumor Segmentation (BRATS) challenge, and they found that GMM and HMRF ranked among the best performing methods [115]. Due to the absence of training data, unsupervised methods rely more strongly on the incorporation of prior knowledge. As gliomas exhibit a wide variability in terms of size, shape, location and appearance, imposing proper prior knowledge is difficult. Fletcher-Heath *et al.* applied knowledge-based FCM to segment non-enhancing gliomas [73]. Clusters containing tumor as well as CSF were separated from white and grey matter based on proton-density (PD-)enhancement. The tumor region was then separated from the ventricular CSF by assuming symmetry of the ventricles across the hemispheres. Finally, extraventricular CSF and other false positives are removed from the tumor mask using morphological post-processing. Kazerooni *et al.* applied FCM combined with manually initiated region growing to segment GBMs on a multi-parametric MRI dataset combining T_2, DWI and PWI [119]. The active tumor region was found by applying FCM to the CBV map, assuming hyperperfusion throughout the region. Saha *et al.* presented a computationally efficient tumor detection algorithm, by assuming symmetry across the hemispheres in healthy brain [218]. A bounding box is obtained within one hemisphere which maximizes the histogram-based dissimilarity with the corresponding contralateral region based on the Bhattacharya coefficient. Many unsupervised segmentation studies incorporate prior knowledge in the form of simplifying rules, limiting their applicability to a subset of gliomas with specific characteristics: the tumor is assumed to be enhancing or non-enhancing, deformation of the surrounding healthy tissue is assumed to be small, multi-focal lesions can not be detected, the lesion does not extend to the contralateral hemisphere, etc. Some methods are limited to a particular set of MRI parameters, as they rely on one or more specific imaging biomarkers.

1.3.3 Segmentation using advanced MRI modalities

The vast majority of segmentation methods have so far only considered cMRI data. Over the past decade, advanced MRI modalities, such as PWI, DWI and MRSI have gained interest in the clinical field, and their added value regarding brain tumor diagnosis, treatment planning and follow-up has been recognized (see Section 1.2). Numerous recent studies have suggested that tumor characterization might benefit from the additional structural, biological and biochemical information provided by these advanced MRI modalities [14, 39, 63, 187]. The RANO working group, who has updated the criteria for response assessment in neuro-oncology in 2010, also acknowledged that inclusion of advanced MRI modalities might aid in monitoring tumor growth [272]. Thus far, only few studies have combined different MRI modalities for glioma segmentation. Di Costanzo *et al.* used DWI, PWI and ^{1}H-MRSI to classify a

set of manually delineated regions of interest in the non-necrotic part of the tumor and the peritumoral region [64]. Step-wise linear discriminant analysis was applied to the multi-parametric (MP-)MRI data to differentiate between regions of gross tumor, edema, tumor/edema, tumor infiltration and normal tissue. Highest classification accuracy was reported when all MRI modalities were included. Verma *et al.* applied support vector machines to a combined set of cMRI and DTI parameters to segment intensity-based tissue profiles on a voxel-wise basis [259]. They provide tissue probability maps as well as hard segmentation of enhancing tumor, non-enhancing tumor and edema in the pathological region. Zikic *et al.* applied a decision forest classifier on a set of voxel-wise intensities and context-aware features from cMRI and DTI [292]. They report an improvement of the segmentation accuracy when DTI is included compared to using only cMRI data. Kazerooni *et al.* are among the first to explore unsupervised classification methods for MP-MRI based glioma segmentation [119]. They employed spatial fuzzy C-means clustering to a combination of cMRI, DWI and PWI data, followed by a region growing step.

1.4 Aims of the thesis

Automated volumetric segmentation of the tumor and its subcompartments would be of immense added value in the management and follow-up of brain tumor patients. This thesis aims at developing (semi-)automated methods for brain tumor segmentation and for characterizing tumor heterogeneity using MP-MRI data. The proposed methods are based on an unsupervised learning technique called Non-negative matrix factorization (NMF). NMF provides an additive parts-based model, thereby assessing the relative proportion of the different tissue types on a voxel-wise basis. Nowadays, most state-of-the-art segmentation algorithms are based on a supervised classification approach. Such methods require extensive training datasets with a uniform acquisition protocol, to account for the wide heterogeneity in tumor appearance among glioma grades. Being an unsupervised method, NMF can directly be applied to any MP-MRI dataset of any tumor patient. This thesis explores whether NMF-based methods can be competitive with state-of-the-art by incorporating prior knowledge into the segmentation algorithm.

A hierarchical NMF variant (hNMF) is proposed, which essentially consists of 2 levels of NMF. Computational efficiency of the hNMF algorithm is improved by considering advanced initialization methods, such as the successive projection algorithm. hNMF performance in terms of brain tumor segmentation is explored and compared to other common unsupervised classifiers, including several single-level NMF and clustering methods. The hierarchical structure allows hNMF

to better differentiate tissue types with similar MP-MRI signatures, such as non-enhancing tumor and edema.

So far, the vast majority of brain tumor segmentation algorithms have only considered cMRI data. However, as advanced MRI modalities, such as PWI, DWI and MRSI have gained interest in the clinical field, numerous recent studies have suggested that tumor characterization might benefit from the additional structural, biological and biochemical information provided by these advanced MRI modalities. Therefore, this thesis will also investigate whether segmentation performance can be increased by including advanced MRI modalities into the MP-MRI dataset. Based on the hNMF algorithm, the added value of the individual MRI modalities (cMRI, PWI, DWI and MRSI) will be explored.

As gliomas exhibit a wide variability in terms of size, shape, location and appearance, imposing proper prior knowledge is difficult. Many unsupervised segmentation algorithms in literature are limited to a subset of gliomas or to a specific MP-MRI feature set, due to the inclusion of restrictive prior knowledge. We aimed at developing a semi-automated NMF-based segmentation framework that is applicable to all types of gliomas and to any MP-MRI dataset, by incorporating user-defined voxel selection in the pathological tissue regions. Regularization terms are considered to promote spatial consistency and sparseness of the tissue abundance maps. A morphological post-processing procedure exploits the location of the selected voxels along with tissue adjacency constraints to further enhance tumor segmentation. Robustness of the segmentation results against user input variability will be thoroughly verified, as it is a key aspect of any semi-supervised method. Segmentation performance of the semi-automated NMF-based method will be compared against state-of-the-art.

This research has been conducted in close collaboration with the Department of Radiology of UZ Leuven and the Department of Radiology of UZ Gent, and is part of a joint research project which was funded by the Research Foundation - Flanders (FWO grant G.0869.12N). Experience with MRI acquisition techniques at both university hospitals was employed to develop efficient MP-MRI scanning sequences, to assess the added value of the individual MRI modalities in the management of glioma patients. Patients with untreated glioma were recruited for the project and MP-MRI datasets of the patient cohorts were provided by both hospitals. Manual segmentations of the tumor and its subcompartments were provided by Dr. Sofie Van Cauter from UZ Leuven and by Dr. Marjan Acou from UZ Gent. These manual delineations served as ground-truth for evaluating the developed segmentation algorithms.

1.5 Outline of the thesis and personal contributions

Figure 1.13 gives a structured overview of the thesis. Chapter 1 introduces the different MRI modalities, their role in the management of glioma patients, and brain tumor segmentation methods. Chapter 2 discusses the NMF model, which forms the basis of the segmentation methods that are developed throughout the thesis. The MP-MRI datasets from UZ Leuven and UZ Gent, as well as pre-processing methods of the individual MRI modalities are discussed in Chapter 3. The hNMF algorithm is considered in Chapters 4 and 6 for characterizing tumor heterogeneity and for tumor segmentation. Chapter 5 explores different initialization methods for improving computational efficiency of the NMF algorithms. Chapters 7 and 8 evaluate the semi-automated NMF-based segmentation framework on different MP-MRI datasets. Finally, the general conclusions and future perspectives are formulated in Chapter 9.

Chapter 2 introduces the NMF model, which forms the basis of the segmentation methods that are developed throughout the thesis. Several of the most common NMF algorithms are explained. The validation metrics that will be used for evaluating the quality of tissue differentiation and tumor segmentation based on NMF output are also introduced.

Chapter 3 describes the MP-MRI datasets from UZ Leuven and UZ Gent, onto which NMF segmentation methods will be applied throughout the thesis. Acquisition parameters and pre-processing methods for the individual MRI modalities are discussed. One of the main personal contributions in this chapter involved the optimization of the metabolite quantification. Several quantification methods were explored, including QUEST and AQSES. Finally, AQSES-MRSI was used and a common frequency shift was introduced into the algorithm, which improved metabolite quantification and allowed the differentiation of Lip and Lac, and of Gly and mI. Pre-processing and quantification of DWI and PWI metrics, and co-registration of the different modalities was also largely personal work.

In **Chapter 4**, the hNMF algorithm is introduced and applied to the MP-MRI dataset from UZ Leuven. A hierarchical NMF variant has been introduced in previous work for differentiating active tumor and necrosis from healthy brain using MRSI data [153]. In this chapter, the hNMF algorithm has been modified to be able to combine different MRI modalities, and to allow for a more refined tissue differentiation. Also, an iterative voxel selection procedure was developed to minimize cross-contamination between the tissue sources after the first level of NMF. The hNMF algorithm is applied to the full MP-MRI dataset, as well as to reduced MRI datasets to assess the added value of the individual MRI

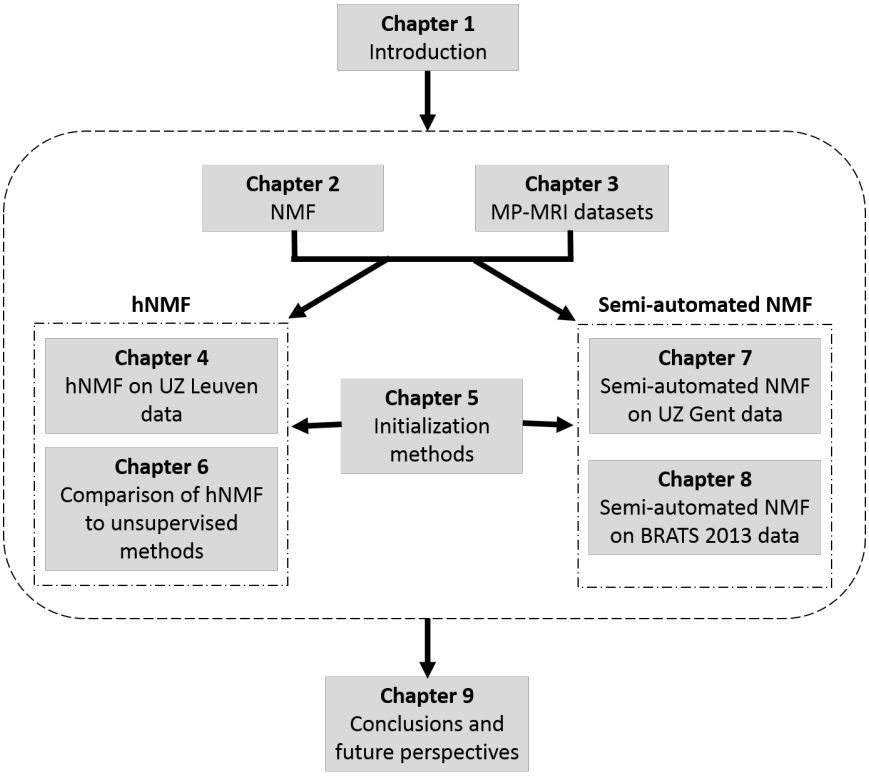

Figure 1.13: Schematic overview of the thesis.

modalities. The work from this chapter was published in a slightly modified version in [221].

Chapter 5 explores the usage of different initialization methods for the non-convex NMF problem. The main novelty in this chapter lies in the consideration of the Successive projection algorithm (SPA) as an initialization method. Whereas SPA is commonly considered as a direct endmember extraction tool for NMF, it is shown that better segmentation performance can be achieved by using SPA as an initialization method. SPA is compared to other common NMF initialization methods in terms of segmentation quality, rate of convergence and computational efficiency. This chapter has been submitted as a jounral paper for peer review as [220]. A partial version of this work has been presented in [221].

Chapter 6 compares hNMF performance in terms of segmentation quality to

several common unsupervised classifiers, including 2 single-level NMF methods, fuzzy c-means clustering, Gaussian mixture modeling and spectral clustering. The comparison is made on both the UZ Leuven and UZ Gent MP-MRI datasets, using several NMF initialization methods. Computational efficiency of the hNMF algorithm has been improved by considering advanced initialization and by replacing the iterative voxel selection procedure after the first level of NMF by a k-means clustering step. This has made hNMF approximately 10 times faster, while maintaining segmentation performance. This work has been published in [224].

In **Chapter 7**, a semi-automated NMF-based segmentation framework is presented based on user-defined voxel selection. L1-regularization is included into the NMF objective function to promote spatial consistency and sparseness of the tissue abundance maps. A morphological post-processing procedure exploits the location of the selected voxels along with tissue adjacency constraints to further enhance tumor segmentation. The method is applied to the UZ Gent MP-MRI dataset. The influence of the regularization terms on the segmentation result is explored. Robustness against user input variability is verified by repeating NMF analyses on the individual patients using different voxel selection. This work has been submitted as a journal paper for peer review as [223].

In **Chapter 8**, the semi-automated NMF-based segmentation framework which was introduced in Chapter 7 is applied to the publicly available BRATS 2013 Leaderboard dataset. This allows direct comparison of our method to state-of-the-art segmentation methods. Bias field correction is applied to the cMRI data as a pre-processing step, to compensate for magnetic field inhomogeneities. Voxel selection by an experienced radiologist is performed several times on each patient and analyses are repeated to quantify intra-user variability of the segmentation results. The work in this chapter has been accepted for publication in the proceedings of the 2016 International Conference on Signal Image Technology & Internet Based Systems [222].

Chapter 9 summarizes the main achievements and general conclusions of the thesis. Opportunities for future research are discussed.

1.6 Conclusion

Conventional MRI has become the imaging modality of choice in the management of glioma patients. Advanced MRI modalities such as PWI, DWI and MRSI have recently gained interest in the clinical field, as they provide biological, structural and biochemical information which is complementary to the anatomical information provided by cMRI. Although the added value of

PWI, DWI and MRSI is unmistakable, it has become clear that no modality in itself is specific enough to answer clinical questions unambiguously. This has led towards multi-parametric MRI (MP-MRI), i.e. the combination of several MRI modalities providing complementary information. MP-MRI will increase specificity in assessing tumor extent, tumor heterogeneity and early response to treatment. Tumor segmentation is an important but challenging aspect of treatment planning and follow-up of glioma patients. Manual segmentation by a radiologist is still the gold standard, although it is a time-consuming and tedious task. Due to the high clinical relevance, there has been an increasing interest in the machine learning community for automated brain tumor segmentation. Clinical practice would benefit from automated volumetric measurements of the tumor and its subcompartments.

Chapter 2

Non-negative matrix factorization and validation metrics

The methods developed throughout this thesis for brain tumor segmentation and characterization are based on Non-negative matrix factorization (NMF). NMF is able to reveal the basic components in a dataset based on an additive parts-based model. The basic concept of NMF, as well as several of the most common NMF algorithms will be discussed in this chapter. In the last section, several validation metrics are introduced for evaluating the quality of tissue differentiation and tumor segmentation based on NMF output.

2.1 Non-negative matrix factorization

2.1.1 NMF model

Non-negative matrix factorization (NMF) was introduced by Paatero and Tapper in 1994 [186]. It became popular after a publication by Lee and Seung in 1999, in which they proposed a simple multiplicative update algorithm for finding a meaningful parts-based representation of facial images (consisting of the eyebrows, mouth, nose, etc.) [148]. Since then, NMF has become a widely used tool for multi-variate data analysis in the fields of blind source separation and pattern recognition. It has found applicability in many domains, including text

mining [231], image processing [148], hyperspectral imaging [83], computational biology [62], and also brain tumor differentiation based on MRSI data [153, 182].

NMF decomposes a non-negative input matrix X into the product of 2 non-negative factor matrices W and H, providing a low-rank (rank r) approximation:

$$X \approx WH \quad \text{with} \quad X \in R_+^{m \times n}, W \in R_+^{m \times r} \quad \text{and} \quad H \in R_+^{r \times n} \tag{2.1}$$

The NMF model requires the rank r as an input. NMF provides an additive parts-based representation of the input data. The columns of X represent observations (or data points) in m-dimensional space. NMF models each column of X as a weighted sum of the r columns of W. It is thereby assumed that all observations can be adequately modeled by a set of r basic components. These basic components, i.e. the columns of W, are also called the NMF sources. The second factor matrix H contains the weights, or so-called abundances, of the NMF model. Each column of H corresponds to one column of X, and contains the weights of the r sources for that data point. The NMF model comes naturally to many applications, as observations are often perceived as a sum of meaningful basic components. As such, the NMF model provides physically interpretable factors, which is one of its main advantages over other low-rank approximation methods.

A similarity measure is required in order to quantify the difference between X and its approximation WH. The most commonly used similarity measure is the Frobenius norm. It will also serve as the objective function for solving the optimization problem:

$$\min_{W,H} f(W,H) = \min_{W,H} \quad \frac{1}{2}\|X - WH\|_F^2, \quad \forall i,j : W_{i,j} \geq 0, \quad H_{i,j} \geq 0 \tag{2.2}$$

where $\|\|_F$ represents the Frobenius norm. Other similarity measures have been considered as well, such as the Kullback-Leibler divergence. The reader is referred to [267] for a more elaborate discussion on NMF objective functions.

2.1.2 NMF algorithms

One of the main challenges of NMF is to design fast and efficient algorithms generating non-negative factors (W, H) that minimize the objective function. Despite its attractive features, i.e. sparse low-rank representation of the input data with interpretable factors, NMF is an NP-hard problem with non-unique solutions [256]. As finding a globally optimal solution within reasonable computation time is infeasible, practical algorithms aim at finding locally optimal solutions. Many algorithms are iterative, and exploit the fact that NMF reduces to an efficiently solvable convex non-negative least squares problem

when fixing one of the factor matrices while optimizing the other one [82]. In fact, most NMF algorithms adhere to the following framework:

(0) Initialize factor matrices by W_0 and H_0

For k=0,1,2,... do:

(1) Fix H_k and find W_{k+1} such that $\|X - W_{k+1}H_k\|_F^2 < \|X - W_kH_k\|_F^2$

(2) Fix W_{k+1} and find H_{k+1} such that $\|X - W_{k+1}H_{k+1}\|_F^2 < \|X - W_{k+1}H_k\|_F^2$

Some of the main NMF algorithms, namely multiplicative updates, hierarchical alternating least squares and convex NMF are discussed below. The reader is referred to literature for a discussion of other common algorithms, such as projected gradient descent and quasi-Newton methods [51].

Multiplicative updates

Lee and Seung proposed multiplicative update (MU) rules to effectively minimize the Frobenius norm as defined in Equation (2.2) [148]. They showed that the Frobenius norm is non-increasing when using MU. The algorithm is defined as follows:

Algorithm 1: Multiplicative updates

Input: $X \in R_+^{m \times n}$, initial factor matrices $W_0 \in R_+^{m \times r}, H_0 \in R_+^{r \times n}$
While stopping criterion not reached **do**:
$W \leftarrow W \circ \frac{XH^T}{W(HH^T)}$;
$H \leftarrow H \circ \frac{W^TX}{(W^TW)H}$;
end while

In the MU rules of Algorithm 1, $\circ$ represents an element-wise multiplication and the ratio represents an element-wise division. The MU algorithm converges relatively slowly, especially when dealing with dense matrices [49]. Secondly, the multiplicative nature of the update rules implies that any zero element will remain zero throughout the algorithm. A simple but efficient solution to this issue, is to replace zero values by a very small positive value [81].

Hierarchical alternating least squares

Looking at the previously formulated general framework for iteratively updating W and H, an obvious approach would be to directly apply non-negative least squares fitting in an alternating way to both convex subproblems. In fact, such an implementation has been formulated [120]. However, this computation, which can be performed using active-set methods, is relatively costly. Alternatively, one might consider an algorithm computing approximative solutions to the convex subproblems, thereby reducing the cost per iteration and obtaining an inexact but faster optimization scheme. An obvious approach would be to apply alternating least squares (ALS) without non-negativity to W and H. Negative elements appearing in W and H during iteration can be brought back to the positive orthant by replacing them by zero or a small positive value [81]. However, standard ALS has been shown to suffer from unstable convergence properties and sensitivity to noise [51]. Hierarchical alternating least squares (HALS) was first introduced by Cichocki *et al.* [50]. HALS successively updates each column of W/row of H with an optimal and easy to compute closed form solution. This is possible because each row of W only affects the corresponding row of WH. Hence, elements of W belonging to different rows can be separately optimized. The HALS algorithm is formulated as follows:

Algorithm 2: Hierarchical alternating least squares

Input: $X \in R_+^{m \times n}$, initial factor matrices $W_0 \in R_+^{m \times r}, H_0 \in R_+^{r \times n}$

While stopping criterion not reached **do**:

Compute $A = XH^T$ and $B = HH^T$.

for $k = 1 : r$ **do**

$$W_{:k} \leftarrow \max\left(\epsilon, \frac{A_{:k} - \sum_{l \neq k}^{r} W_{:l} B_{lk}}{B_{kk}}\right);$$

end for

Compute $C = W^T X$ and $D = W^T W$.

for $k = 1 : r$ **do**

$$H_{k:} \leftarrow \max\left(\epsilon, \frac{C_{k:} - \sum_{l \neq k}^{r} D_{kl} H_{l:}}{D_{kk}}\right);$$

end for

end while

Convex NMF

Convex NMF was first proposed by Ding *et al.* [66]. It assumes that the sources (i.e. the columns of W) are convex combinations of the columns of X.

In other words, the sources are required to lie within the column space of X. A non-negative auxiliary matrix A is introduced to impose this constraint:

$$W = XA \quad \text{such that} \quad X \approx XAH \tag{2.3}$$

Multiplicative update rules are defined to alternatingly revise A and H towards convergence. Convex NMF comes with the appealing property that without any additional constraints, A and H are both sparse. Although Convex NMF can also handle mixed-sign data, its updating rules are simplified when considering purely non-negative data (the ratios in the updating rules represent the Hadamard quotient):

Algorithm 3: Convex NMF

Input: $X \in R_+^{m \times n}$, initial factor matrices $H_0 \in R_+^{r \times n}$, $A_0 \in R_+^{n \times r}$

While stopping criterion not reached **do**:

$$H^T \leftarrow H^T \sqrt{\frac{(X^T X)A}{H^T A^T (X^T X)A}};$$

$$A \leftarrow A \sqrt{\frac{(X^T X)H^T}{(X^T X)AHH^T}};$$

end while

2.2 Validation metrics

2.2.1 Dice score

The Dice score is one of the most commonly used metrics to quantify segmentation overlap between 2 methods:

$$Dice_{\text{tissue}} = 2 \times \frac{A_{\text{tissue,method1}} \cap A_{\text{tissue,method2}}}{A_{\text{tissue,method1}} + A_{\text{tissue,method2}}} \tag{2.4}$$

where $A_{\text{tissue,method1}}$ denotes the area of a certain tissue region as obtained by method1 and $A_{\text{tissue,method2}}$ is the segmented region of the same tissue obtained by method2. The Dice score is expressed as a percentage, where 0% indicates no overlap and 100% represents perfect overlap between both segmentations. The Dice score has been found to be the most suitable metric in assessing spatial alignment between different segmentations [294].

2.2.2 Positive predictive value

The positive predictive value (PPV) is commonly used in diagnostic tests to represent the proportion of positive results that are truly positive:

$$PPV = \frac{TP}{TP + FP} \tag{2.5}$$

where TP represents the number of true positives and FP the number of false positives. In terms of tissue segmentation, PPV values can be calculated for each tissue type, using manual segmentation by an expert rater as the ground-truth.

2.2.3 Sensitivity

Sensitivity (also called the recall) measures the proportion of positive results that are correctly identified as such:

$$Sensitivity = \frac{TP}{TP + FN} \tag{2.6}$$

where TP represents the number of true positives and FN the number of false negatives. In terms of tissue segmentation, sensitivity values can be calculated for each tissue type, using manual segmentation by an expert rater as the ground-truth.

2.2.4 Cosine similarity

Cosine similarity is a measure of similarity between two non-zero vectors x and y that measures the cosine of the angle between them:

$$Cos = \frac{x \cdot y}{\| x \| \| y \|} \tag{2.7}$$

The cosine similarity is calculated as the dot product of the normalized vectors. It is a measure of the angular similarity between 2 vectors, varying between 0 (orthogonal) and 1 (collinear).

2.3 Conclusion

In this chapter, the NMF model has been introduced. NMF provides an additive parts-based representation of the input data. As such, NMF will allow us to perform voxel-wise tissue differentiation when applied to (multi-parametric) MRI data of the brain. Several common NMF implementations that are applied throughout the thesis are discussed. In the final section, validation metrics that will be used for evaluating the quality of tissue segmentation have been introduced.

Chapter 3

Multi-parametric MRI datasets and pre-processing

One of the main objectives of the thesis is to assess up to what extent advanced MRI modalities can aid in the tissue differentiation within gliomas. In this chapter, 2 multi-parametric MRI (MP-MRI) datasets are presented, both of them including cMRI, PWI, DWI and MRSI data. The advanced MRI modalities provide us with biological, structural and biochemical information which is complementary to the anatomical information provided by cMRI. The 2 MP-MRI datasets were acquired at the university hospital of Leuven (UZ Leuven) and at the university hospital of Gent (UZ Gent). They were acquired using a different acquisition protocol, resulting in a different set of MP-MRI features. Both protocols aim at extracting important information within a relatively short scanning time (less than one hour). The main differences between both protocols involve the MRSI and DWI data. Short-TE 2D ^{1}H-MRSI data were acquired at UZ Leuven, whereas long-TE ^{1}H-MRSI data were obtained at UZ Gent in a 3D volume of interest. Regarding diffusion-weighted imaging, an extensive DKI protocol was included at UZ Leuven, allowing the computation of several advanced DTI and DKI related features. At UZ Gent, a basic DWI sequence was used. For each MRI modality, state-of-the-art pre-processing methods are presented to quantify the MP-MRI features from the raw acquired imaging data. Throughout the thesis, both MP-MRI datasets will be analysed under a unifying framework, allowing to explore the added value of the individual MRI modalities in terms of tumor segmentation and characterizing tumor heterogeneity.

3.1 UZ Leuven dataset

3.1.1 Patient population

A total of 24 glioma patients were included in the UZ Leuven dataset (9 females/15 males; age range: 22-78 years; median age: 55 years). None of the patients had neurological disorders other than primary neoplasm and no patients with recurrent disease were included. The lesions were classified according to grade using the 2007 WHO classification [157], with histopathological confirmation in all cases. Out of the 24 patients, 10 were diagnosed with an LGG: 1 patient with a grade I oligodendroglioma, 4 patients with a grade II astrocytoma, 3 patients with a grade II oligodendroglioma, 1 patient with a grade II oligoastrocytoma and 1 with a grade II lesion with mixed findings of pilocytic astrocytoma and oligodendroglioma. The other 14 patients were diagnosed with an HGG: 1 patient with a grade II astrocytoma with focal progression to grade III, 1 patient with grade III astrocytoma, 1 patient with grade III oligoastrocytoma and 11 patients with GBM. All studies were performed according to the guidelines of the UZ Leuven Medical Ethical Committee and informed consent was obtained from all patients before scanning. MP-MRI acquisition was performed on a 3T MR system (Philips Achieva, Best, The Netherlands), using a body coil for transmission and an 8-channel head coil for signal reception. The total scanning time per patient was approximately 45 minutes.

3.1.2 cMRI

cMRI consisted of T_2-weighted imaging, T_1-weighted imaging with contrast enhancement (T1c) and FLAIR. An axial spin echo T_2-weighted MRI was acquired with the following parameters: TR/TE: 3000/80ms; slice/gap: 4/1mm; turbo factor: 10; field of view (FOV): $230 \times 184 mm^2$; acquisition matrix: 400×300. A T_1-weighted 3D spoiled gradient echo MRI scan with contrast administration (0.1mmol/kg body weight meglumine gadoterate (Dotarem, Guerbet, France)) was performed with the following parameters: TR/TE/inversion time (TI): 9.7/4.6/900ms; flip angle: 8°; turbo field echo factor: 180; acquisition voxel size: $0.98 \times 0.98 \times 1 mm^3$; 118 contiguous partitions. An axial FLAIR MRI scan was acquired with the following parameters: TR/TE/TI: 11000/120/2800ms, slice/gap: 4/1 mm, FOV: $230 \times 184 mm^2$, acquisition matrix: 240×134.

3.1.3 PWI

PWI was obtained using Dynamic susceptibility-weighted contrast-enhanced MRI (DSC-MRI) with a gradient-echo EPI sequence. A series of 60 multi-section scans was acquired at 1.35 second intervals, using the following parameters: TR/TE: 1350/30ms; section thickness/gap: 3/0mm; FOV: $200 \times 200mm^2$; acquisition matrix: 112×109; EPI data were acquired during the first pass following a rapid injection of a 0.1mmol/kg body weight bolus of meglumine gadoterate (Dotarem, Guerbet, France) via a mechanical pump at a rate of 4ml/s, followed by a 20ml bolus of saline.

As TR is well below 2s, slice-timing correction of the raw PWI images was not required [236]. Cerebral blood volume (CBV) is by far the most commonly used PWI metric to study brain tumors [40]. Relative CBV (rCBV) maps were computed in ImageJ [226] using the DSCoMAN plugin (Daniel P. Barboriak Lab, Duke University, North Carolina, USA). DSCoMAN computes rCBV based on the method proposed by Boxerman *et al.* [20], compensating for contrast agent leakage due to disruption of the blood-brain barrier (see Section 1.2.3). The measured relaxivity change in each voxel is approximated as a linear combination of the whole-brain average relaxivity change in non-enhancing voxels, $\overline{\Delta R_2^*(t)}$, and its time integral:

$$\Delta R_2^*(t) \approx K_1 \overline{\Delta R_2^*(t)} - K_2 \int_0^t \overline{\Delta R_2^*(\tau)} d\tau \tag{3.1}$$

where $\Delta R_2^*(t)$ represents the measured relaxivity change in a pixel, the first term $K_1 \overline{\Delta R_2^*(t)}$ reflects the uncontaminated relaxivity change and the second (K_2-)term reflects the effects of leakage. Linear least squares fitting is applied to compute the weighting factors K_1 and K_2 over all the voxels. Corrected relaxivity curves are obtained by only withholding the K_1-term. A positive value of K_2 indicates that the leakage effect is dominated by the T_2^* effect, whereas a negative value indicates that the T_1 effect is dominant. This model may be more effective in correcting for leakage effects than gamma-variate fitting functions, which tend to underestimate rCBV [233]. Cerebral blood flow (CBF) maps are not considered, as they correlate strongly with CBV and their computation is not well-conditioned [34].

3.1.4 DKI

As for diffusion-weighted imaging, an extensive diffusion kurtosis imaging (DKI) protocol was included at UZ Leuven. A spin echo EPI DWI sequence was used to acquire the DKI data. DKI gradient settings were optimized as proposed in [196]: implemented b-values were 0, 700, 1000, and $2800 s/mm^2$, applied in respectively 1, 25, 40, and 75 uniformly distributed directions. The following parameters were used throughout the DKI acquisition sequence: TR/TE: 3200/90ms; gradient duration (δ): 20ms; diffusion time interval (Δ): 48.3ms; FOV: $240 \times 240 mm^2$; acquisition matrix: 96×96; number of signal averages: 1; section thickness/gap: 2.5/0mm; parallel imaging: SENSE with factor 2 in the antero-posterior direction. The scan time was approximately 18 minutes.

Occasional venetian blind artefacts in the b_{2800} images were corrected for by applying a slice-wise co-registration to the b_0 reference. EDDY in FSL [112] was used to correct for eddy currents, induced by the rapidly switching magnetic field gradients, and head motion. EDDY is based on registering the individually acquired volumes to a model free prediction of what each volume should look like [1]. It only assumes that closely related gradient directions will produce similar MR signals and opposing gradients will produce identical MR signals. The prediction is calculated using a Gaussian process for which the hyperparameters are estimated directly from the data. To account for Gibbs ringing artefacts, isotropic smoothing was applied using a Gaussian kernel (FWHM=3mm). This smoothing was only applied to the b_0 and b_{700} images, as Gibbs ringing is most prominent at high SNR (i.e. low b-values), and also not to alter the nature of the noise distribution in the highly diffusion-weighted images in order to avoid a noise bias in the subsequent tensor estimation. The DKI model as defined in Equation (1.8) can be structured into a linear regression form based on the natural logarithm of the raw DWI signals [11]. The 22 unknown DKI parameters were estimated using a constrained weighted linear least squares (WLLS) model [258]. The weights are the squares of the predicted MR signals, as reconstructed from a previous estimate using unweighted linear least squares fitting. Constraints were imposed on the diffusion and kurtosis parameters according to [123, 247]. Unlike other advanced diffusion parameter estimators, the WLLS model does not require noise estimation and it is unbiased under mild conditions [258]. Mean diffusivity (MD), fractional anisotropy (FA) and mean kurtosis (MK) maps were computed from the DTI and DKI tensors as described in [196]. Finally, the diffusion maps were non-rigidly co-registered to a T2 reference using ExploreDTI [149], to account for EPI-induced geometrical distortions due to susceptibility variations [108].

3.1.5 MRSI

A 2D ^{1}H-MRSI sequence with short TE was included in the UZ Leuven acquisition protocol. The 2D MRSI grid was manually positioned in the central axial plane of the tumor based on axial T2 images. Point-resolved spectroscopy (PRESS) was used as the volume selection technique [18] with a bandwidth of 1.3kHz for the slice-selective pulses. Acquisition parameters were as follows: TR/TE: 2000/35ms; FOV: $160 \times 160 mm^2$; maximal region of interest: $80 \times 80 mm^2$; section thickness: 10mm; acquisition voxel size: $10 \times 10 mm^2$; reconstruction voxel size: $5 \times 5 mm^2$; receiver bandwidth: 2000Hz; samples: 2048; number of signal averages: 1; water suppression method: multiple optimizations insensitive suppression train (MOIST) [180]; first- and second-order pencil beam shimming; parallel imaging: sensitivity encoding (SENSE) with reduction factors of 2 (left-right) and 1.8 (anterior-posterior). Scan time per patient was 3 minutes and 30 seconds. Automated pre-scanning optimized the shim in order to yield a peak width consistently under 20Hz full-width half-maximum.

The raw MRSI data were exported from the Philips platform after standard post-processing (zero-filling in k-space, transformation from k-space to time domain, automatic phase correction and eddy current correction). Further processing was done with the in-house SPID software [198] in MatLab 2012b (MathWorks, Massachussets, USA). Initially, 2 metabolite quantification methods were considered: QUEST [208] and AQSES [199]. Both QUEST and AQSES are time domain methods that make use of a metabolite basis set to fit the MRSI signals. More specifically, each MRSI signal is modeled as a linear combination of K damped, phased and frequency-shifted metabolite templates:

$$\hat{y}(t) = \sum_{k=1}^{K} a_k \exp(j\phi_k) \exp(-d_k t + 2\pi j f_k t) v_k(t) + b(t) + w(t) \tag{3.2}$$

where $\hat{y}(t)$ represents the modeled time signal, a_k are the amplitudes of metabolite profiles from the basis set, ϕ_k and d_k are phase shifts and Lorentzian damping factors that account for the necessary corrections to the basis set profiles. $b(t)$ is the baseline signal, which accounts for macromolecular components and possibly other highly damped signals that are not covered by the basis set. $w(t)$ represents the water component. Model-based time domain quantification methods such as QUEST and AQSES can be expressed as a non-linear least squares problem of fitting the model $\hat{y}(t)$ to a measured signal:

$$\min_{a_k, \phi_k, d_k, f_k} \sum_{t=0}^{N} | \, y(t) - \hat{y}(t) \, |^2 \tag{3.3}$$

where N represents the number of time samples of the measured signal. We used an in-vitro measured short-TE metabolite basis set for quantification. Metabolite solutions were measured under constant environmental conditions (temperature=37°C, pH=7.0). The following metabolites were initially quantified: (Lip1+Lac), NAA, Glx, Cre, Cho, mI.

QUEST[1] uses a Levenberg-Marquardt algorithm to minimize the objective function as defined in Equation (3.3), but it does not directly incorporate the baseline and water component into the signal model (i.e. $b(t) = 0$ and $w(t) = 0$ in Equation (3.2)). Water filtering needs to be performed as a pre-processing step. Hankel-Lanczos singular value decomposition with partial re-orthogonalization (HLSVD-PRO) was used, as it has been shown to outperform other SVD-based water filtering techniques [136]. QUEST deals with the baseline signal prior to quantification, using a truncation procedure: as the baseline components are highly damped, they are assumed to be removed from the time signal by omitting a limited number of the first time points. Based on an initial quantification from the truncated signal, the first time points can be reconstructed using back-extrapolation and a baseline-free time signal can be derived for quantification.

AQSES is implemented using a variant of the variable projection (VARPRO) algorithm. Linear parameters from the signal model are projected out of the least squares problem. The remaining non-linear variables are obtained via an iterative minimization algorithm of the Levenberg-Marquardt type. Maximum-phase FIR filtering (MPFIR) is used throughout the iterative quantification procedure to model the water signal [200]. AQSES allows to model the baseline using a basis set of splines. A regularization term can be added to the objective function to promote smoothness of the baseline. We did not include such a baseline model, as it further complicates quantification, increasing the degrees of freedom and so the risk of converging to a local minimum. Instead, MPFIR has been shown to suppress highly damped components, thereby increasing the signal-to-baseline ratio [200]. Hence, a considerable reduction of the baseline components could be obtained with MPFIR.

More robust quantification results were obtained with AQSES compared to QUEST. Water filtering was more reliable with MPFIR than with HLSVD-PRO, especially in the case of improper water suppression in the raw MRS signal or low-quality spectral data. This is in accordance with a previous study comparing MPFIR and HLSVD-PRO on short-TE ^{1}H-MRS data [200]. Furthermore, baseline removal through truncation as used by QUEST was not found to be robust in all cases. Quantification results were sometimes poor for highly damped metabolites, mainly Glx and mI, as their main signal also occurs

[1]QUEST implementation in SPID was used, which differs from the more popularized implementation in jMRUI.

in the first (truncated) time points and little metabolite signal remains after truncation. Direct comparison of spectral fitting between QUEST and AQSES is difficult, as MPFIR introduces a frequency-dependent phase shift to the time signal [200]. Figure 3.1 compares the Cre maps obtained with QUEST and AQSES for a GBM patient with relatively poor spectral data quality. QUEST results in a non-uniform map with large local variations in Cre level. AQSES provides a smooth map with increased Cre levels in the healthy brain, and a clear reduction in Cre within the GBM lesion. It must be noted that comparison between QUEST and AQSES was based solely on visual inspection of spectral fitting and the metabolite maps, due to the lack of an undisputed objective metric for assessing fitting quality of the individual metabolites.

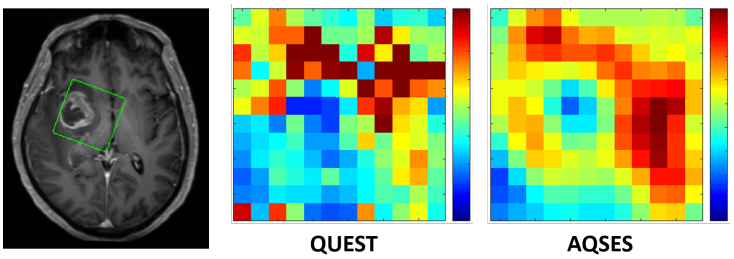

Figure 3.1: Cre maps obtained using QUEST and AQSES for a GBM patient with low-quality spectral data. The MRSI region of interest is marked in green on T1c on the left.

AQSES-MRSI [57], an extension of AQSES, was considered to further improve quantification results. Whereas QUEST and AQSES apply metabolite fitting to each voxel individually, AQSES-MRSI includes prior knowledge in the form of proximity constraints on the spectral parameters within an MRSI grid and adds a penalty term to the objective function that promotes smooth spectral parameter maps. Although MPFIR was able to largely reduce the baseline signal, quantification was still influenced by residual broad components due to the lack of an actual baseline model. To avoid such errors, more stringent constraints were applied to the frequency variations of the individual metabolites. The main cause for frequency shifts in MRS(I) spectra is related to variations of the static magnetic field B_0, causing changes in the Larmor frequency. These field variations may occur in the form of local B_0 inhomogeneity (e.g. due to imperfect shimming or large susceptibility variations) or as a temporal field drift [69]. This kind of frequency shift affects the entire spectrum, thereby shifting all metabolites in the same way. Instead of allowing a relatively large frequency variation to each individual metabolite, we allowed for a common frequency shift of the entire spectrum, and only a small individual shift per metabolite. Such

kind of constraint is also applicable in other spectral quantification tools such as LCModel [204]. The maximal common frequency shift was set to 50rad/s, whereas the maximal individual shift per metabolite was set to 10rad/s. The maximal damping variation per metabolite was set to 30rad/s.

Another concern is the combined quantification of different metabolites, such as the lipid peak at 1.3ppm and the lactate doublet. Both metabolites are related to different biological phenomena: whereas lipids are a hallmark of necrosis and occur mainly in GBMs, lactate is related to hypoxia and occurs also in lower grade gliomas [30]. Nevertheless, they are commonly quantified together, due to their strong spectral overlap and a lack of phase difference at short TE. Spectral editing techniques, which require the acquisition of 2 different TE per phase-encoding step, are used to reliably differentiate between Lip1 and Lac [39]. Another less recognized metabolite pair that is often combined is the mI multiplet and the Gly resonance at 3.55ppm. mI is considered to be essential for cell growth, acting as an osmosis regulator [155]. Its signal is clearly present in healthy brain, and elevated mI levels have been reported in LGGs. Gly, on the other hand, mainly occurs in HGGs and its signal is near the noise threshold in healthy brain [48]. Since the separation of these strongly overlapping metabolites could be of high relevance, we attempted to differentiate between Lip1 and Lac, and between mI and Gly using the tight frequency constraints as explained in the previous paragraph. Proper metabolite profiles were added to the basis set for Lac and Gly. It was hypothesized that the individual overlapping metabolite profiles would not interfere with each other, as the individual frequency shifts had been strongly restricted. Although the proposed procedure is probably not as reliable as spectral editing, we did systematically find intermediate levels of Lac and Lip1 in grade III gliomas, and only strong elevation of the Lip1 signal in GBMs, as expected. Elevated levels of Gly, which was previously quantified in combination with mI, were found in many GBM patients (see Figure 3.2).

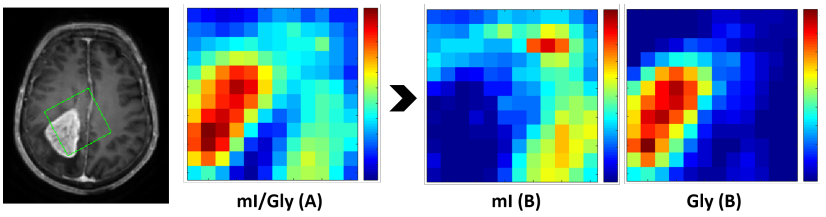

Figure 3.2: Combined (A) and separate (B) quantification of mI and Gly in a GBM patient. Increased Gly and reduced mI (B) are common in GBM.

3.1.6 MP-MRI dataset

In summary, quantified MP-MRI features for the UZ Leuven dataset are: T2, T1c, FLAIR, rCBV, MD, FA, MK, Lip1, Lac, NAA, Glx, Cre, Cho, mI and Gly. Figure 3.3 shows a co-registered set of MP-MRI maps of a GBM patient.

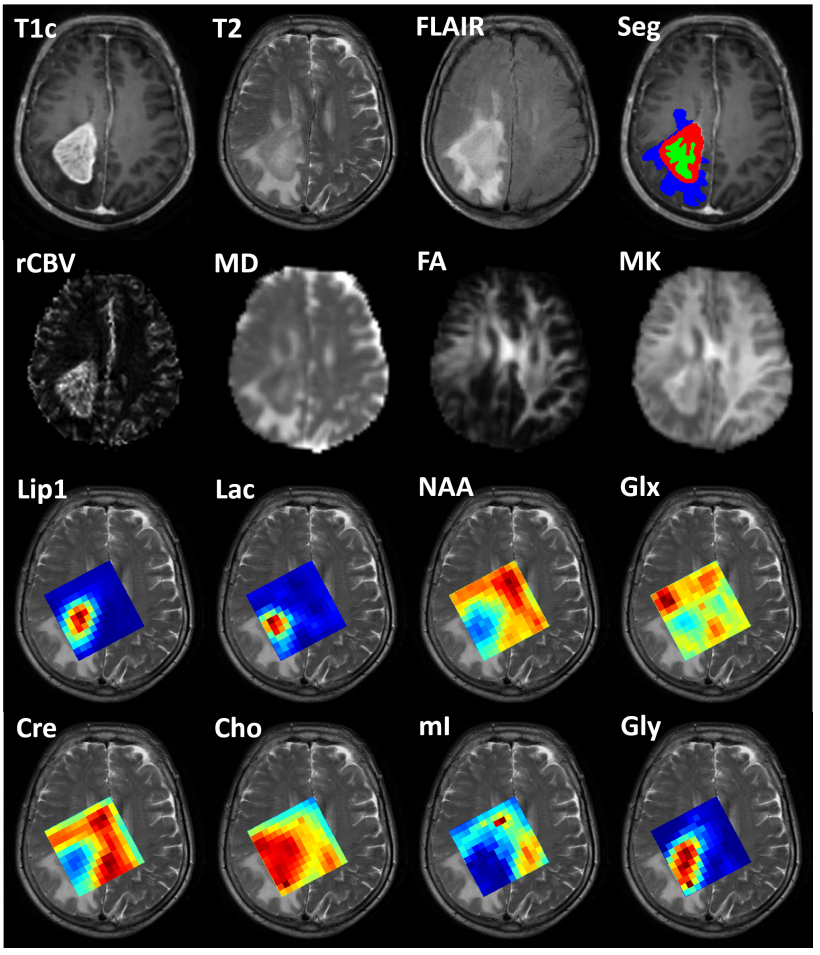

Figure 3.3: Co-registered set of MP-MRI maps of a GBM patient from UZ Leuven. 'Seg' indicates manual segmentation of active tumor (red), necrosis (green) and edema (blue). Color scaling differs among the different metabolites.

3.2 UZ Gent dataset

3.2.1 Patient population

The UZ Gent dataset consisted of 21 patients who were diagnosed with an HGG (7 females/14 males; age range: 25-79 years; median age: 44 years). The lesions were classified according to grade using the 2007 WHO classification [157]. 1 patient was diagnosed with a grade II astrocytoma with focal progression to grade III, 1 patient had a grade II oligodendroglioma with focal progression to grade III, 1 patient had a grade II oligo-astrocytoma with focal progression to grade III, 6 patients were diagnosed with anaplastic oligo-astrocytoma, 2 patients had an anaplastic astrocytoma with focal progression to GBM and 10 patients were diagnosed with GBM. None of the patients had neurological disorders other than primary neoplasm and no patients with recurrent disease were included. The Ghent University Hospital local ethics committee allowed a retrospective analysis of the data. All patients underwent an MP-MRI acquisition protocol consisting of cMRI, PWI, DWI and MRSI. MR examinations were performed on a 3T Siemens Trio Tim scanner (Erlangen, Germany), using a standard 12-channel phased array head coil.

3.2.2 cMRI

cMRI included T_1-weighted imaging with and without contrast enhancement, and FLAIR imaging. T_1-weighted imaging consisted of a 3-dimensional gradient-echo sequence (MPRAGE) with isotropic voxels. 176 sagittal slices were acquired using the following scanner settings: TR/TE/TI: 1550/2.39/900ms; FOV: $220 \times 220 mm^2$; voxel size $0.9 \times 0.9 \times 0.9\text{mm}^3$; acquisition matrix: 256×256; parallel imaging with GRAPPA factor 2; flip angle: 9°. The same MPRAGE sequence was repeated at the end of the scanning protocol after administration of gadolinium contrast. FLAIR consisted of a 3-dimensional T_2-weighted inversion recovery sequence with isotropic voxels. 176 sagittal slices were acquired using the following scanner settings: TR/TE/TI: 6000/421/2100ms; FOV: $250 \times 235 mm^2$; voxel size $1.0 \times 1.0 \times 1.0\text{mm}^3$, acquisition matrix: 256×238; parallel imaging with GRAPPA factor 2.

3.2.3 PWI

Perfusion-weighted imaging was performed by using a lipid-suppressed, T2*-weighted echo-planar imaging DSC-MRI sequence with the following parameters: TR/TE: 1000/29ms; FOV: $230 \times 230\text{mm}^2$; slice thickness: 5.0mm; acquisition matrix: 128×128; in-plane voxel size: $1.8 \times 1.8\text{mm}^2$; 15 axial slices, distance factor: 30%; parallel imaging with GRAPPA factor 2; flip angle: 90°. A series of 90 multi-section volumes was acquired at 1 second intervals. The first 10 acquisitions were performed before contrast agent injection to establish a pre-contrast baseline. At the tenth acquisition, a 0.1mmol/kg body weight bolus of gadobutrol (Gadovist, Bayer, Germany) was injected with a power injector (Spectris, Medrad Inc., Pennsylvania, USA) at a rate of 4ml/s through a 18-gauge intravenous catheter, immediately followed by a 20 ml bolus of sodium chloride solution at 4ml/s.

Slice time correction was not required since TR was well below 2s [236]. As for the UZ Leuven dataset, rCBV maps were calculated using the DSCoMAN plugin in ImageJ. DSCoMAN computes rCBV based on the method proposed by Boxerman *et al.* [20], compensating for contrast agent leakage due to disruption of the blood-brain barrier. For a more detailed discussion of the method, see Section 3.1.3.

3.2.4 DWI

The UZ Gent protocol included a DWI acquisition consisting of a fast single-shot gradient-echo echo-planar imaging sequence. b_0 images were acquired, serving as a non-diffusion-weighted reference. In addition, b-values of 500 and $1000 s/mm^2$ were measured in 3 orthogonal directions. Each measurement consisted of 30 axial slices using the following scanner settings: TR/TE: 5400/80ms; FOV: $264 \times 264\text{mm}^2$; voxel size $2.0 \times 2.0 \times 3.0\text{mm}^3$; parallel imaging with GRAPPA factor 2; number of averages = 3. The scan time was 2 minutes and 10 seconds.

To compensate for eddy current distortion and head motion, each volumetric scan was co-registered to the b_0 reference using an affine transformation. Based on the mono-exponential model as described in Equation 1.6, a weighted linear least squares method [258] was used to compute ADC values in 3 orthogonal directions using the b_0, b_{500} and b_{1000} images. MD values were obtained from averaging ADC values over the 3 directions. Since cMRI did not include T_2-weighted imaging, the raw b_0 images were also included in the MP-MRI feature set, serving as a T_2-weighted reference.

3.2.5 MRSI

A 3D ^{1}H-MRSI sequence with long TE was included in the acquisition protocol. The volume of interest, consisting of 2 adjacent MRSI slabs, was positioned on axial reconstructions of the 3D FLAIR sequence to include tumor, peri-lesional edema and normal brain tissue. A Stimulated Echo Acquisition Mode (STEAM) pulse sequence was used for 3D spatial localization [170]. The volume was completely enclosed within the brain and 8 outer volume suppression slabs were placed outside the volume of interest. Magnetic resonance parameters were: TR/TE: 1700/135ms; FOV: $160 \times 160\text{mm}^2$; region of interest: $80 \times 80\text{mm}^2$; voxel size: $10 \times 10 \times 15\text{mm}^3$; flip angle: 90°; acquisition bandwidth: 1200Hz; number of samples: 1024; number of averages: 3. Weak water suppression was used. Automatic shimming with manual fine-tuning of the B_0 magnetic field was used as well as iterative semi-automatic optimization of the transmitter voltage. The scan time was 11 minutes and 45 seconds.

Standard post-processing of the raw MRSI signals consisted of zero-filling in k-space, transformation from the k-space to the time domain and automatic phase correction and eddy current correction. Further processing was done with the in-house SPID software [198] in MatLab 2012b (MathWorks, Massachussets, USA). AQSES-MRSI, which promotes spatial smoothness of the spectral parameters, was used for metabolite quantification [57]. MPFIR was used for suppressing the residual water signal, as it had been shown to outperform other water suppression techniques on the UZ Leuven MRSI data (see Section 3.1.5). Inclusion of a baseline model was not required as the highly damped macromolecular components have been attenuated at long TE. Constraints applied to the frequency variation of the individual metabolite profiles were similar to those applied on the UZ Leuven data. A common frequency shift was first determined by a rough initial fitting of the MRSI signals using only the NAA, Cre and Cho metabolite profiles. In a next step, all metabolites were fitted and small deviations from the common frequency shift were allowed to the individual metabolite profiles. The maximal additional frequency shift per metabolite was set to 10rad/s. The maximal damping variation per metabolite was set to 30rad/s. The following metabolites were quantified: Lip1, Lac, NAA, Glx, Cre and Cho.

3.2.6 MP-MRI dataset

In summary, quantified MP-MRI features for the UZ Gent dataset are: T1, T1c, FLAIR, rCBV, MD, b0, Lip1, Lac, NAA, Glx, Cre and Cho. Figure 3.4 shows a co-registered set of MP-MRI maps of a GBM patient.

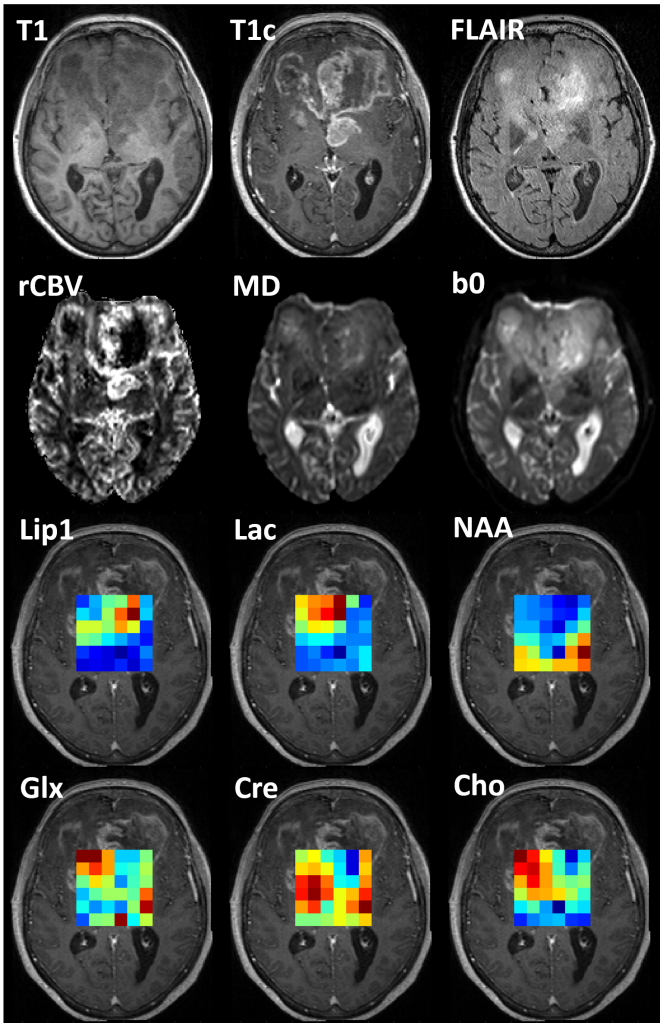

Figure 3.4: Co-registered set of MP-MRI maps of a GBM patient from the UZ Gent dataset. Color scaling differs among the different metabolites.

Chapter 4

Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using MP-MRI data

The aim of this chapter is to present a hierarchical variant of non-negative matrix factorization to characterize tumor heterogeneity in glioma patients and to differentiate abundances of active tumor, necrosis and edema on a voxel-wise basis. Hierarchical NMF (hNMF) has been proposed in a previous study, to differentiate active tumor and necrosis from healthy brain tissue by using MRSI data [153]. In this chapter, the hNMF algorithm will be adapted to be able to deal with MP-MRI data and to provide a more refined tissue characterization. hNMF will be applied to the MP-MRI dataset of UZ Leuven, including cMRI, PWI, DKI and MRSI. The added value of the individual MRI modalities on the tissue characterization will be explored. The work presented in this chapter is based on [225].

4.1 Introduction

Conventional magnetic resonance imaging (cMRI) has become the imaging modality of choice upon suspicion of a brain tumor. cMRI allows for localization of the lesion and assessment of radiological features like contrast enhancement and presence of necrosis, which might aid in establishing differential diagnosis. However, in a significant number of cases, cMRI is unable to provide reliable diagnosis and grading [90], which is crucial for optimal treatment planning and prognosis. Gliomas, and in particular HGGs are known to present a high degree of heterogeneity, both inter- as well as intra-patient [190, 109]. Spatial differences in cellular phenotype and malignancy are common, and pathological tissue types are known to merge into each other. Due to this heterogeneity, there is a risk of underdiagnosing when biopsy sampling is required. Intra-tumoral heterogeneity is not adequately reflected in cMRI [109], making it an improper tool for biopsy targeting.

Over the past decade, advanced MRI modalities such as PWI, DWI and MRSI have gained interest in the clinical field, and their added value regarding brain tumor diagnosis, treatment planning and follow-up has been recognized [39, 283]. Many studies have investigated the potential of individual MRI modalities to characterize tumor grade, detect full tumor extent and assess early success of therapy. Although the added value of PWI, DWI and MRSI is unmistakable, it has become clear that no modality in itself is specific enough to answer clinical questions unambiguously. This conclusion has led towards multi-parametric MRI (MP-MRI), i.e. the combination of several MRI modalities to provide complementary information on the structural, hemodynamic and/or biochemical characteristics of the tissue. Several studies have shown that MP-MRI improves diagnostic accuracy when it comes to grading gliomas compared to individual MRI modalities [230, 252, 293]. Other studies have reported improved differentiation of intra-tumoral and peri-tumoral tissue types by applying supervised classification methods to MP-MRI data [64, 114, 259].

Although the spatial heterogeneity of HGGs is partially explained by the presence of different stages of the disease throughout the lesion [190], it is also attributed to the co-existence of several intra-tumoral tissue types within the same region, such as the presence of focal necrosis within viable tumor tissue [156], tumor infiltration in peri-tumoral edema [137] and the formation of vasogenic edema in the viable tumor region [174]. Therefore, characterization of the tumoral and peri-tumoral tissue might benefit from a model that incorporates the concept of tissue mixtures. Blind source separation techniques have been applied to brain tumor MRSI data for this purpose. In particular, non-negative matrix factorization (NMF) methods have been shown to give most promising results [182, 219]. NMF aims to extract physically meaningful sources, corresponding

to tissue-specific patterns. NMF assesses the relative contribution of each tissue type within each voxel, assuming the dataset can be modeled as a linear combination of the constituent tissue types. In this study, a hierarchical NMF technique is presented and applied to an MP-MRI dataset combining cMRI, PWI, DKI and ^{1}H-MRSI to characterize tumor heterogeneity, providing a comprehensive representation of the spatial distribution of tissue types in the tumor region. Limited user interaction is required for specifying the number of tissue types present in the region of interest. Although voxels of mixed tissue contributions are allowed, the method benefits from the assumption that a large number of voxels will contain mainly one tissue type, given the relatively high resolution of cMRI, at which all MRI maps are resampled. Tissue-specific patterns are obtained on a patient-by-patient basis. The usage of MP-MRI data provides a detailed tissue characterization and delimitation at the voxel level, and is able to improve source separation of the different tissue types. To assess the added value of individual MRI modalities, hNMF is also applied to reduced MP-MRI datasets, leaving out one modality at a time, and to cMRI data only.

4.2 MP-MRI dataset

The institutional human ethics review board approved this prospective study. Twenty-four patients with confirmed diagnosis of glioma underwent an MP-MRI examination at UZ Leuven. Fourteen out of the 24 patients were diagnosed with HGG, 3 of which had a grade III glioma and 11 had a GBM (grade IV). The other 10 patients suffered from LGG, with 1 patient having a grade I oligodendroglioma and all others a grade II glioma. MRI acquisition was performed on a 3T MR system (Philips Achieva, The Netherlands), using a body coil for transmission and an 8-channel head coil for signal reception. Table 4.1 gives an overview of the acquisition protocol and the obtained MP-MRI parameters for each MRI sequence. A detailed description of the acquisition protocol and data pre-processing methods for the UZ Leuven dataset can be found in Section 3.1.

To be able to perform voxel-wise analyses, all MP-MRI maps were co-registered and brought to the same spatial resolution. The T1c images were resliced such that one slice coincided with the central plane of the MRSI slab. This resliced T1c set served as a reference. Skull-stripping was applied to all modalities prior to co-registration. T2, FLAIR and rCBV maps were rigidly co-registered to the T1c reference set using SPM8 (Wellcome Department of Imaging Neuroscience, University College London, UK). The normalized mutual information criterion was used for co-registration [161], with cubic spline interpolation for reslicing. The diffusion-weighted metrics (MD, FA and MK) were non-rigidly co-registered

	UZ Leuven dataset
cMRI	**T2, T1c, FLAIR**
PWI	*Gradient echo DSC-MRI* **rCBV**
DKI	*Spin echo DWI (b_0, b_{700}, b_{1000}, b_{2800} in resp. 1, 25, 40, 75 directions)* **MD, MK, FA**
MRSI	*2D short TE ^{1}H-MRSI* **Lac, Lip1, NAA, Glx, Cre, Cho, Gly, mI**

Table 4.1: Schematic overview of the MP-MRI protocols and derived parameters of the UZ Leuven dataset.

to T2 using ExploreDTI [149], to account for geometrical distortions induced by the EPI DKI sequence. Co-registration was not possible for the MRSI data, due to their low spatial resolution and the limited 2D region of interest. The metabolite maps were spatially aligned with the T1c reference, assuming no patient motion had occurred between MRSI and T1c acquisition. T1c was used as a reference because it was acquired close to MRSI within the scanning protocol. A band of MRSI voxels at the outer edges of the region of interest was omitted to avoid chemical shift displacement artefacts and lipid contamination (band proportional to the grid size, maximally 3 voxels wide for a 16×16 grid). All MP-MRI maps were upsampled to the T1c spatial resolution of $0.98 \times 0.98 \times 1mm^3$ using cubic spline interpolation. On average, approximately 3200 voxels were included in the region of interest per patient.

A total of 15 original MP-MRI parameters were included in the feature set: T2, T1c, FLAIR, rCBV, MD, FA, MK, Lip1, Lac, NAA, Glx, Cre, Cho, mI and Gly. Additional intensity-based features were added to the feature set to include localized spatial information. An in-plane local neighborhood of 3×3 and 5×5 voxels was used to calculate average intensity features that were assigned to the central voxel. These smoothed features were computed for the 7 non-metabolic MP-MRI maps and added to the feature set. They were not included for the metabolic features because of their low spatial resolution. This resulted in a total of 29 MP-MRI features (i.e. 7 non-metabolic features, 14 smoothed non-metabolic features and 8 metabolic features). The metabolic features were rescaled by dividing them by the maximum metabolite value throughout the MRSI region of interest, such that relative amplitudes were unchanged between the different metabolites. For the non-metabolic features, each feature's full range was rescaled to [0-1].

4.3 Hierarchical non-negative matrix factorization

The concept of non-negative matrix factorization (NMF) has been introduced in Section 2.1. NMF decomposes a non-negative input matrix X into the product of 2 low-rank non-negative factor matrices W and H:

$$X \approx WH \quad \text{with} \quad X \in R_+^{m \times n}, W \in R_+^{m \times r} \quad \text{and} \quad H \in R_+^{r \times n} \tag{4.1}$$

The rank r of the factor matrices is required as an input to the NMF algorithm. The n columns of X represent the data points, each of which contains m features. NMF approximates each data point as a weighted sum of r source vectors. The source vectors are the columns of W and the weights (or abundances) per data point are the columns of H. The goal of our work is to use NMF to characterize tumor heterogeneity based on MP-MRI data. The MP-MRI features belonging to one voxel are stacked into a vector. All these vectors are grouped as the columns of the input matrix X. Each voxel's feature vector will be modeled as a weighted sum of the r sources in W. The sources are expected to represent the different tissue types, so each source contains a tissue-specific signature of MP-MRI features. The columns of H contain the weights of the r sources for the corresponding columns of X. One row of H contains the weights for one tissue type over all the voxels. Such a row can be spatially transformed back to the imaging domain, to obtain tissue-specific abundance maps. As we are dealing with image intensities, and tissue abundance within a voxel can not be negative, the non-negativity constraints apply naturally to the input data as well as to the factor matrices. The most commonly used similarity metric to quantify the difference between X and its NMF approximation WH is based on the squared Euclidean distance:

$$\min_{W,H} f(W,H) = \min_{W,H} \quad \frac{1}{2}\|X - WH\|_F^2 \quad ,\forall i,j : W_{i,j} \geq 0, \quad H_{i,j} \geq 0 \tag{4.2}$$

Throughout this chapter, we will consider accelerated hierarchical alternating least squares (aHALS) to solve the basic NMF problem as defined in Equation (4.2). aHALS belongs to the family of HALS NMF algorithms (see Section 2.1.2). It benefits from a computationally efficient implementation and fast convergence compared to other NMF methods [82].

Hierarchical non-negative matrix factorization (hNMF) consists of 3 steps: 2 levels of NMF and a re-combining step. The individual steps are discussed in detail below. Figure 4.2 gives a schematic overview of the hNMF procedure.

Step1: A rank-2 NMF analysis is performed on all voxels within the region of interest (ROI). This results into 2 sources: $W_{:,1}$ and $W_{:,2}$, and their corresponding abundances, $H_{1,:}$ and $H_{2,:}$. It is assumed that the pathological tissue types (viable tumor, and possibly necrosis and/or edema) and the normal tissue types (white matter, gray matter, cerebro-spinal fluid (CSF), blood vessels) present within the ROI will be divided over these 2 sources. As a result, a voxel with a dominant tissue type will have a higher abundance for the source that represents that tissue type. As NMF suffers from scaling ambiguity, each abundance map was rescaled such that the maximum abundance value over all the voxels was set to 1. The sources were rescaled accordingly.

Step2: After step1, we assume that each tissue type is represented by either $W_{:,1}$ or by $W_{:,2}$. All voxels have 2 abundance values, one for each source. Based on these abundance values, the voxels will be assigned to one of the two sources. An iterative voxel selection criterion is applied, based on voxel-wise abundance ratio maps calculated over the whole ROI:

$$H_{ratio1}(i) = \frac{H_{1,:}(i)}{H_{1,:}(i) + H_{2,:}(i)} \qquad \text{and} \tag{4.3}$$

$$H_{ratio2}(i) = \frac{H_{2,:}(i)}{H_{1,:}(i) + H_{2,:}(i)} \tag{4.4}$$

where $H_{1,:}(i)$ represents the abundance of voxel i for $W_{:,1}$ and $H_{2,:}(i)$ is the abundance for $W_{:,2}$. We then consider a variable threshold t to perform the voxel assignment. At each fixed value of t, we assign those voxels with H_{ratio1} larger than t to $W_{:,1}$ and voxels with H_{ratio2} larger than t to $W_{:,2}$. We let t range from 0.5 to 1, with t=0.5 corresponding to voxels with equal abundances for both sources and t=1 corresponding to voxels with a nonzero abundance for only one of the 2 sources. For each value of t, we apply NMF to the group of voxels assigned to $W_{:,1}$, and to the group of voxels assigned to $W_{:,2}$. To determine the rank for both NMF analyses, the number of tissue types represented by each source needs to be determined. Based on visual inspection of the cMRI data, the number of tissue types present within the ROI is determined. Each tissue type is then assigned to one of the 2 sources by considering which source shows the highest abundances in the corresponding tissue region.

As t increases, only voxels that have a more clear dominant source are withheld. In this way, tissue types represented by $W_{:,1}$ do not have to fit voxels that contain significant contributions of tissue types represented by $W_{:,2}$ and vice-versa, minimizing cross-contamination between the tissue sources. Figure 4.1

illustrates how only the voxels with a more dominant abundance are withheld as t increases. The t-range was divided into 20 steps, as a finer step size did not significantly improve the results any further.

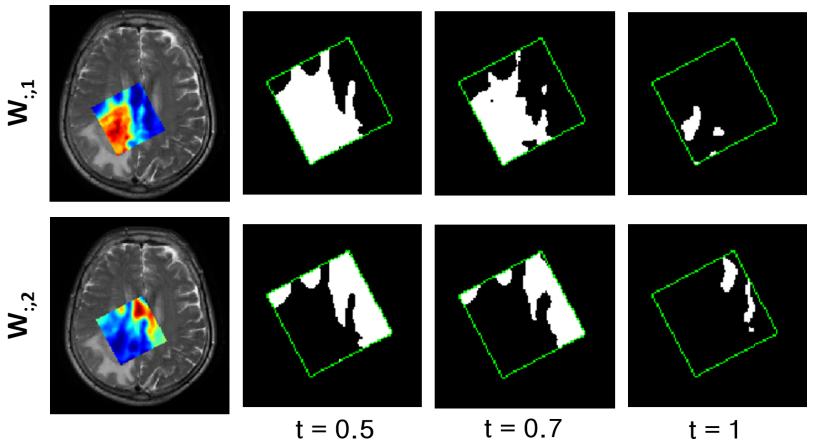

Figure 4.1: Left: abundances maps of the 2 sources after hNMF step1. Voxel selection for the second level of NMF is shown for t=0.5, t=0.7 and t=1.

Step3: The sources obtained from the NMF analyses in step2 represent tissue-specific signatures. For each value of t, the tissue sources obtained from both regions are re-combined to calculate the abundances over the whole ROI using non-negative least-squares fitting (NNLS). To automatically determine the optimal threshold value t, a criterion that promotes sparsity to the voxel-wise tissue abundances is introduced. The optimal threshold t then corresponds to the highest H^t_{ratio3} value, where:

$$H^t_{ratio3} = \sum_i \frac{H^t_{max}(i)}{\sum_j H^t_j(i)} \tag{4.5}$$

where $H^t_j(i)$ denotes the abundance of source j in voxel i using threshold t, and $H^t_{max}(i)$ represents the highest abundance in voxel i at threshold t. The summation over i covers the whole ROI and the sum in the denominator goes over all the j sources. A higher value of H^t_{ratio3} can be interpreted as a more sparse solution: on average, the voxels will have a more dominant tissue type. The final hNMF factorization is the one belonging to the optimal threshold t.

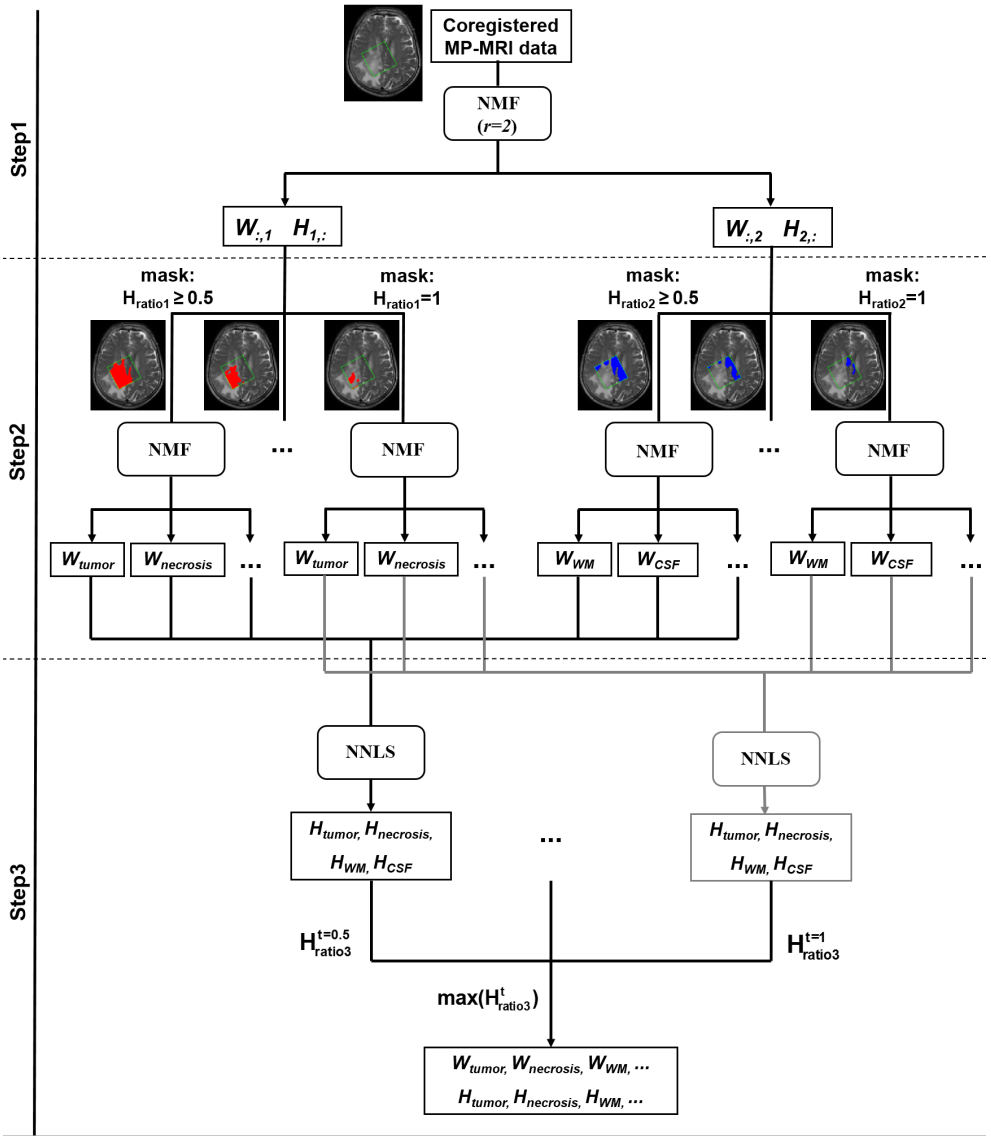

Figure 4.2: Schematic overview of the 3 steps of the hNMF procedure.

hNMF analyses

hNMF analysis was applied to each patient's MP-MRI dataset. The analyses were performed on the MRSI (2D) ROI, as only voxels within this region had a full MP-MRI dataset. It must be noted that NMF is a non-convex optimization problem, meaning that the NMF result will depend on the initialization of the factor matrices. To obtain robust and fairly reproducible results, the hNMF procedure was repeated 20 times for each patient, with random initialization of the elements of W and H.

Besides exploring the potential of hNMF in characterizing tumor heterogeneity using the MP-MRI data, we also wanted to assess the added value of the individual MRI modalities. To this purpose, hNMF was not only applied to the full MP-MRI dataset but also to 5 reduced MRI datasets: 4 reduced MP-MRI datasets by omitting consecutively the cMRI, PWI, DKI and MRSI data from the full MP-MRI dataset, and one dataset considering only the cMRI data.

hNMF validation

Validation of the hNMF results was performed in 2 ways: by means of the pathological tissue abundance maps and by means of the pathological tissue sources. In both cases, validation relied on the manual segmentation of the tumor and its subcompartments by an experienced neuro-radiologist (SVC, 6 years of experience in brain tumor research).

For each patient, a tissue segmentation was obtained from the abundance maps, assigning each voxel to 1 tissue class. This was done by applying k-means clustering to the abundance values, with k set to the number of tissue types resulting from hNMF. Each cluster was initialized with one abundance value set to 1 and all others set to 0. The segmentation obtained from hNMF was compared to the manual segmentation by means of the Dice score (see Section 2.2.1):

$$Dice_{\text{tissue}} = 2 \times \frac{A_{\text{tissue,hNMF}} \cap A_{\text{tissue,manual}}}{A_{\text{tissue,hNMF}} + A_{\text{tissue,manual}}} \tag{4.6}$$

where $A_{\text{tissue,hNMF}}$ denotes the area of a certain tissue region as defined by hNMF and $A_{\text{tissue,manual}}$ is the segmented region of the same tissue as defined by the radiologist. For comparability of the results, Dice scores are reported for pathological tissue classes similar to those defined in the Multi-modal Brain Tumor Segmentation (BRATS) challenge [168]: active tumor (viable tumor), the tumor core (viable tumor + necrosis) and the whole tumor region (viable tumor + necrosis + edema).

The validation of the sources is based on the hypothesis that the sources correspond to actual tissue types. The average feature vector was calculated over all voxels belonging to the same tissue type, based on the radiologist's segmentation. As a measure of correlation, the cosine similarity (see Section 2.2.4) is calculated between the averaged feature vector and the corresponding source obtained from hNMF. Correlation values were calculated for active tumor, necrosis and edema.

The mean, standard deviation and range of the Dice scores and correlation coefficients are reported for the HGG and LGG group of patients. To assess whether improved tissue differentiation was statistically significant for the full MP-MRI dataset compared to the reduced (MP-)MRI datasets, the non-parametric one-tailed Wilcoxon signed rank test was applied. A statistical significance level of $p<0.05$ was considered.

4.4 Results

Figure 4.3 illustrates some of the relevant MRI features as well as the hNMF results for patient 16 (GBM) using the full MP-MRI dataset. Five tissue types were differentiated within the region of interest: white matter, CSF, active tumor, necrosis and edema. hNMF results are only shown for the 3 pathological tissue types. Looking at the tissue sources, active tumor is mainly differentiated from the other pathological tissue types based on elevated Cho, T1c and rCBV. These are common characteristics of malignant tumor, and in particular of GBM: higher Cho reflects increased cellular proliferation, elevated T1c shows disruption of the blood-brain barrier in the tumorous vessels and elevated rCBV represents the high vascularization of the active tumor region. The necrotic region is mainly characterized by low metabolite levels except for elevated levels of Lip and Lac, which are hallmarks of respectively necrosis and hypoxia. Elevated levels of FLAIR are seen in necrosis and edema due to their high (non-fluidic) proton density. Reduced levels of MK and high levels of FA, probably indicating the presence of white matter structures, distinguish edema from the other sources.

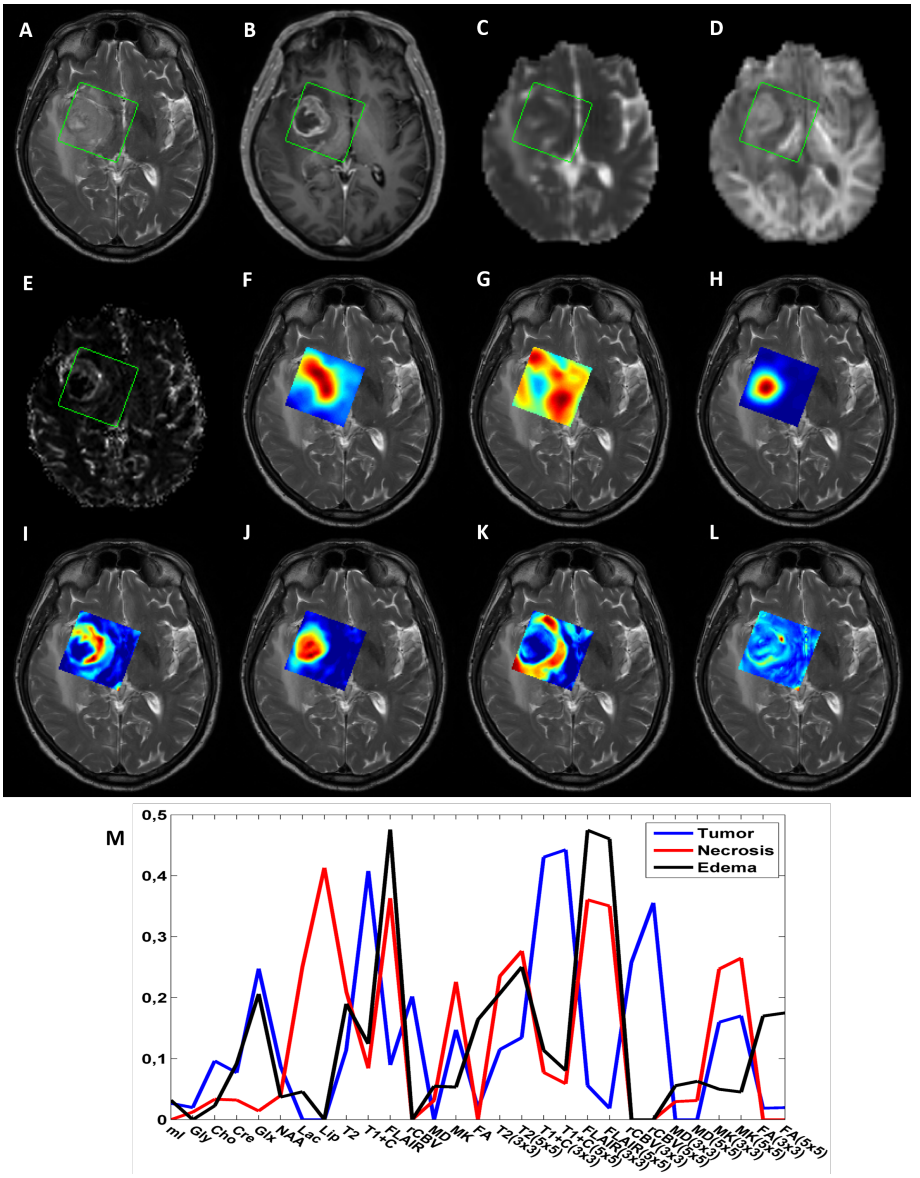

Figure 4.3: MP-MRI maps of different modalities and hNMF results for GBM patient 16. The shown input maps are: T2 (A), T1c (B), MD (C), MK (D), rCBV (E), Cho (F), Cre (G) and Lip (H). The green box indicates the MRSI ROI. hNMF abundance maps are shown for tumor (I), necrosis (J) and edema (K). The residual error map is shown in (L) and the pathological hNMF sources in (M).

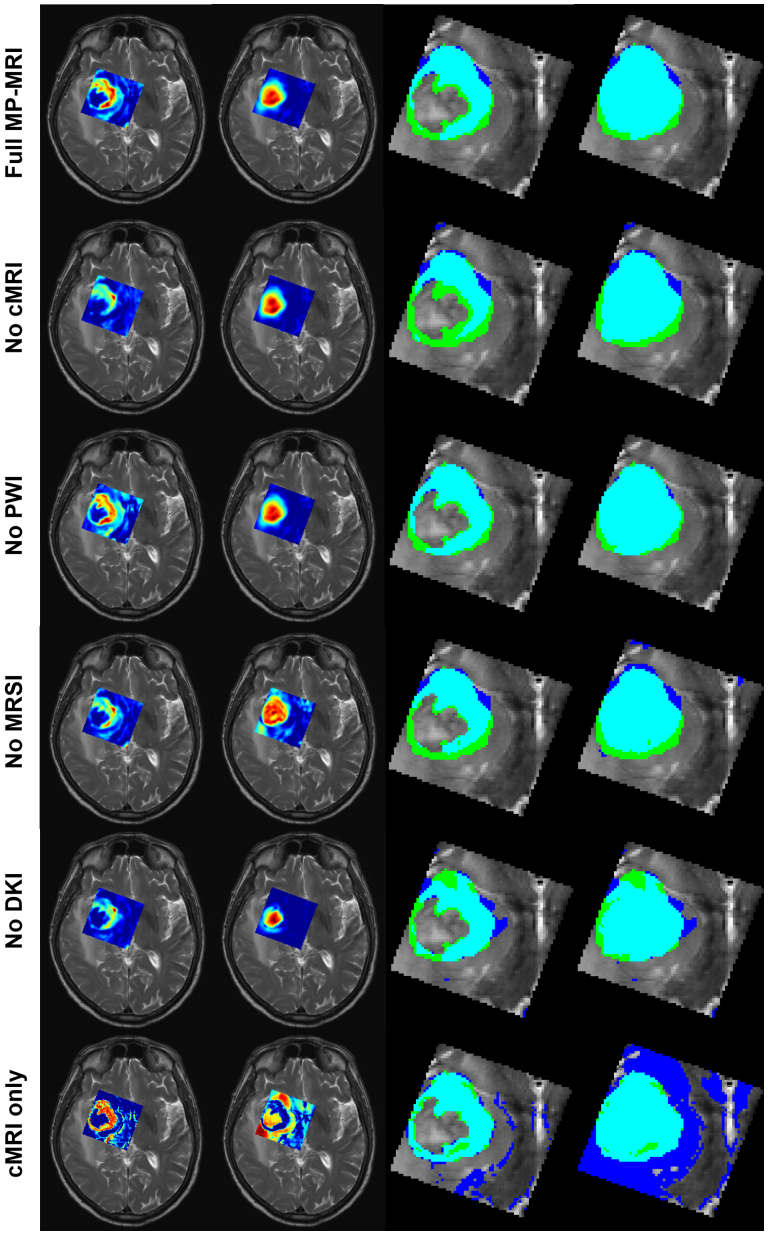

Figure 4.4: Comparison of the hNMF results for the different MRI datasets for GBM patient 16. hNMF abundance maps are shown for active tumor and necrosis. A comparison of the segmentation obtained with hNMF (blue) and by the radiologist (green) is shown for active tumor and for the tumor core. Cyan indicates segmentation overlap.

Figure 4.4 compares the abundance maps of active tumor and necrosis for patient 16 using the full and the 5 reduced MRI datasets. The resulting segmentations for active tumor and the tumor core are shown on the right. Best segmentation results were obtained with the full MP-MRI dataset and the no-PWI dataset (i.e. MP-MRI without PWI), both giving similar abundance maps and hNMF segmentations. For the no-cMRI case, tumor segmentation was reduced, missing a larger portion of the manual delineation. For no-MRSI, some voxels outside the pathological region were assigned to the tumor core. In the case of no-DKI, a small portion of edema adjacent to the tumor region was labelled as tumor. Performance was clearly lowest when considering cMRI only, where necrosis could no longer be differentiated from edema and voxels outside the pathological region were assigned to the tumor core. The Dice scores and correlation coefficients of the individual patients for all considered MRI datasets can be found in the Appendix A at the end of the thesis.

4.4.1 Dice scores

Table 4.2 reports the mean, standard deviation and range of the Dice scores for active tumor, the tumor core and the whole tumor region of the HGG patients for all considered datasets. For the full MP-MRI dataset, mean Dice scores of 78%, 85% and 83% were found for active tumor, tumor core and whole tumor, respectively. Lower mean Dice scores were found for all reduced MRI datasets for the 3 tissue classes. Statistical significance could not be shown in a few cases: for the no-PWI case for active tumor and the tumor core, and for the no-DKI case for active tumor. Considering the 4 reduced MP-MRI datasets, the lowest mean Dice scores are found for no-cMRI. Especially for active tumor, a decrease of 13% was found compared to full MP-MRI. Besides that, the decrease in mean Dice score ranged between 2% and 8% for the 3 tissue classes for the reduced MP-MRI datasets. When considering only cMRI, a larger decrease of the Dice scores is found: 9% for tumor, 14% for tumor core and 15% for whole tumor. The standard deviation is lower and the minimum Dice score is higher for full MP-MRI for each tissue class compared to the reduced datasets.

	Dice score active tumor [%]					Dice score tumor core [%]					Dice score whole tumor [%]				
	All	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**	**All**	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**	**All**	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**
Mean	**78**	**65**	**76**	**76**	**71**	**69**	**85**	**77**	**81**	**79**	**79**	**71**	**83**	**75**	**79**	**77**	**76**	**68**
Standard deviation	10	14	19	12	16	18	11	14	18	14	17	17	14	15	17	14	18	18
Range [Min-Max]	59-93	39-91	20-93	56-92	43-93	38-91	54-98	40-92	36-98	48-92	40-98	39-95	54-98	40-98	36-98	48-98	40-97	39-97
p-value	-	0.001*	0.540	0.180	0.03*	0.004*	-	0.003*	0.280	0.002*	0.01*	0.001*	-	0.002*	0.03*	0.003*	0.001*	0.001*

Table 4.2: Mean, standard deviation and range of the Dice scores for the HGG patients. Results are reported for active tumor, the tumor core and the whole tumor region using the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only. p-values report statistical significance of higher Dice scores for the full MP-MRI dataset (* indicates statistical significance for $p<0.05$).

Table 4.3 shows the Dice scores for the whole tumor region of the LGG patients for all considered datasets. As necrosis does not occur in LGGs and edema was found within the MRSI ROI in only 1 patient, Dice scores are only reported for the whole tumor region. The mean Dice score is 85% for the full MP-MRI dataset and is lower for all reduced MRI datasets. Statistical significance could only be shown compared to no-cMRI, for which the mean Dice score was 73%. For all other reduced datasets, only a slight decrease in mean Dice score was found, ranging from 1% to 4%.

	Dice score whole tumor [%]					
	All	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**
Mean	**85**	**73**	**81**	**84**	**84**	**81**
Standard deviation	14	20	15	16	17	18
Range [Min-Max]	50-96	42-95	52-96	46-96	41-96	38-96
p-value	-	0.003*	0.10	0.32	0.85	0.07

Table 4.3: Mean, standard deviation and range of the Dice scores for the LGG patients. Results are reported for the whole tumor region using the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only. p-values report statistical significance of higher Dice scores for the full MP-MRI dataset (* indicates statistical significance for $p<0.05$).

4.4.2 Correlation coefficients

Table 4.4 reports the mean, standard deviation and range of the correlation coefficients for active tumor, necrosis and edema of the HGG patients for all considered datasets. Correlation coefficients are above 0.90 in nearly all cases, indicating that the tissue source vectors obtained with hNMF are highly representative for the corresponding tissue region. Higher correlation coefficients are found for some reduced MRI datasets compared to full MP-MRI, and statistical significance of higher correlation coefficients for full MP-MRI could not be shown in most cases. Differences in correlation coefficients are relatively small among the MRI datasets: compared to full MP-MRI, variations ranged from a decrease of 0.06 to an increase of 0.05.

	Correlation coefficient tumor					Correlation coefficient necrosis					Correlation coefficient edema				
	All	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**	**All**	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**	**All**	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**
Mean	**0.91**	**0.92**	**0.89**	**0.96**	**0.85**	**0.87**	**0.97**	**0.92**	**0.97**	**0.95**	**0.95**	**0.94**	**0.96**	**0.93**	**0.94**	**0.97**	**0.93**	**0.93**
Standard deviation	0.09	0.07	0.13	0.02	0.16	0.16	0.03	0.13	0.03	0.05	0.06	0.09	0.04	0.06	0.07	0.04	0.07	0.11
Range [Min-Max]	0.71-0.99	0.76-1	0.63-1	0.93-1	0.54-1	0.59-1	0.90-1	0.58-0.99	0.90-0.99	0.86-0.99	0.81-1	0.70-1	0.90-1	0.84-0.99	0.80-1	0.89-1	0.77-1	0.76-0.99
p-value	-	0.52	0.30	0.98	0.12	0.19	-	0.02*	0.26	0.04*	0.33	0.21	-	0.17	0.19	0.95	0.11	0.56

Table 4.4: Mean, standard deviation and range of the correlation coefficients for the HGG patients. Results are reported for active tumor, necrosis and edema using the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI and for cMRI data only. p-values report statistical significance of higher correlation coefficients for the full MP-MRI dataset (* indicates statistical significance for $p<0.05$).

Table 4.5 reports the mean, standard deviation and range of the correlation coefficients for the tumor region of the LGG patients for all considered datasets. The mean correlation coefficient is not lower than 0.90 for any dataset. For the full MP-MRI dataset, a mean correlation coefficient of 0.95 was found. Higher correlation coefficients are obtained for the reduced MRI datasets, except for no-cMRI. The lower correlation coefficient for no-cMRI is probably related to the lower Dice score, indicating a higher mismatch between the hNMF tumor source and the actual tumor region. The difference in the correlation coefficients was not found to be statistically significant in any case.

	Correlation coefficient active tumor					
	All	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**
Mean	**0.95**	**0.90**	**0.96**	**0.98**	**0.98**	**0.97**
Standard deviation	0.08	0.15	0.09	0.02	0.03	0.03
Range [Min-Max]	0.75-1	0.57-1	0.72-1	0.94-1	0.93-1	0.89-1
p-value	-	0.08	0.88	1	0.95	0.81

Table 4.5: Mean, standard deviation and range of the correlation coefficients for the LGG patients. Results are reported for active tumor using the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI and for cMRI data only. p-values report statistical significance of higher correlation coefficients for the full MP-MRI dataset (* indicates statistical significance for $p<0.05$).

4.4.3 Influence of the averaged features

As explained in Section 4.2, spatially averaged non-metabolic features were added to the feature set using a local neighborhood of 3×3 and 5×5 voxels. To assess the effect of adding these averaged features, Table 4.6 compares the mean Dice scores and correlation coefficients obtained with hNMF on the full MP-MRI dataset with and without the averaged features. Mean Dice scores were higher for all tissue classes and for HGGs as well as for LGGs when the averaged features were included. The increase in Dice score ranged between 3% and 6% and and was statistically significant in all cases, based on a one-tailed Wilcoxon signed rank test.

The mean correlation coefficients, on the other hand, remained nearly constant for all pathological tissue types and both in HGGs and LGGs. For the calculation of the correlation coefficient, the average feature vector is calculated over the corresponding tissue region. Since the added features are also averaged (over the local neighborhood of the voxel), it is to be expected that these additional features will show nearly the same correlation as the original features.

	HGGs		**LGGs**	
	With avg feats	**No avg feats**	**With avg feats**	**No avg feats**
Dice active tumor [%]	78*	72	-	-
Dice tumor core [%]	85*	80	-	-
Dice whole tumor [%]	83*	80	85*	81
Corr coeff active tumor	0.91	0.92	0.95	0.95
Corr coeff necrosis	0.97	0.97	-	-
Corr coeff edema	0.96	0.95	-	-

Table 4.6: Comparison of the mean Dice scores and correlation coefficients for hNMF with and without averaged non-metabolic features. Results are reported for HGGs and LGGs separately. * indicates statistical significance of higher values, using the one-tailed Wilcoxon signed rank test with $p<0.05$.

4.5 Discussion

4.5.1 Comparison with previous NMF studies

To the authors' knowledge, this is the first study in which hNMF is applied to multi-parametric MRI data for brain tumor characterization. Several studies have applied NMF to brain tumor MRSI data for tissue differentiation [153, 182, 219]. The concept of hierarchical NMF was introduced in a previous study, reporting improved tissue differentiation for hNMF compared to single-level NMF on simulated as well as in vivo MRSI data of GBM patients [153]. Single-level NMF was initially also considered in the current study, but statistically significantly better results were found for hNMF when applied to the MP-

MRI. A direct comparison of hNMF with several single-level NMF methods will be made in Chapter 6. The advantage of using MP-MRI data compared to only MRSI data is twofold. Firstly, the higher resolution of the added MRI features will allow a more detailed and accurate delineation of the tumor subregions. Previous studies also considered an hNMF-based segmentation for validation purposes. In Li *et al.* [153], this validation was only done visually, hampering comparison of the results. Ortega-Martorell *et al.* [182] did quantify the correspondence between the segmentation obtained from hNMF and post-mortem Ki67 immunostaining applied to 7 GBM bearing mice, resulting in a mean accuracy of 83.7% for the tumor region. However, safe thresholds were applied to the proliferation index obtained from immunostaining for labelling the voxels, such that MRSI voxels near the tumor-normal tissue interface were not always considered. Furthermore, differentiation between tumor and necrosis was not considered. Secondly, the low resolution of the MRSI data results in considerable partial volume effects, with many voxels containing tissue mixtures. Especially in the case of small lesions, there might not be any pure voxels containing only necrosis or tumor, which will make source separation more difficult. Additionally, it was pointed out in [153] that the MRSI spectrum of the tumor source can approximately be modeled as a linear combination of the necrotic and the normal tissue source. This complicates the separability of the tissue types. Adding other MRI features to the dataset will further decorrelate the tissue sources, improving their separability.

The H_{ratio3} criterion that was implemented to find the optimal threshold t in hNMF step3 assumes that most voxels will contain mainly one tissue type. However, when only MRSI data are considered, it is not evident that the same criterion can be applied successfully, given the low spatial resolution of the MRSI data. By applying the optimal threshold t to both the H_{ratio1} and H_{ratio2} map, voxels that contain considerable contributions of both sources after step1 are not considered for determining the tissue-specific patterns in hNMF step2. Since in step2, each voxel is fully modeled by either the tissue types represented by $W_{:,1}$ or the tissue types represented by $W_{:,2}$, these mixed voxels will act as a confounding factor for determining the tissue sources. The optimal threshold value t varied considerably among patients and there was no clear trend. Lower H_{ratio3} values were usually found at the highest t-values, as voxel selection gets so much reduced that one or more tissue types are no longer represented within the voxel selection. In general, the variability of the validation scores as a function of t was higher in patients with a higher number of tissue types within the MRSI ROI. In patients with only 2 or 3 tissue types, Dice scores and correlation coefficients are often rather constant over the t-range, such that selection of the optimal threshold gave little improvement.

The rationale behind adding the averaged non-metabolic features with 3×3 and

5×5 averaging windows is twofold. Firstly, some of the original non-metabolic features showed high intra-tissue variability (e.g. rCBV in the active tumor region), which lead in some cases to a mis-labelling of the 'outlier' voxels. By adding the averaged features, hNMF was better able to cope with this intra-tissue heterogeneity. Secondly, lower Dice scores were commonly seen in patients due to the low spatial accuracy of the metabolic features, which had more weight in the hNMF results when the averaged features were not included. Statistically significantly higher Dice scores were obtained for all tissue classes when including the averaged features, in HGGs as well as in LGGs. Previous NMF studies assumed the number of sources/tissue types to be a priori known [153, 182]. Automatic detection of the number of sources in NMF is challenging, and none of the proposed methods have reached general applicability [267]. A limitation in the current study is that the number of sources had to be set manually for each individual patient. This was done based on visual inspection of the cMRI data by an experienced neuro-radiologist. If the number of sources is set lower than the actual number of tissue types, sources will represent mixtures of tissue types. If the number of sources is set too high, some tissue types will be modeled by more than one source. Future research should therefore focus on reliable automatic assessment of the number of sources.

4.5.2 Reduced MP-MRI datasets and cMRI only

The hNMF procedure was also applied to reduced MP-MRI datasets, leaving out one MRI modality at a time. Compared to the full MP-MRI dataset, mean Dice scores were lower and the standard deviation was higher for all reduced MP-MRI datasets and for each tissue class. For the HGG patients, the decrease in mean Dice score varied between 2% and 8% in all but one case, showing that hNMF still performs well when less MRI modalities are available. A larger decrease in mean Dice score of 13% was only found for tumor for the no-cMRI dataset. This result shows the importance of T1c to differentiate active tumor from necrosis. For the no-PWI case, Dice scores were only significantly higher for the complete tumor region. One might expect rCBV to aid in differentiating tumor, due to an increased (micro)vascularity in the HGG viable tumor region. But compared to other MRI features, rCBV shows a large heterogeneity within the viable tumor tissue, with some areas not being well perfused at all [158]. In some patients, such low-perfused tumor areas were not delineated as tumor when including PWI, leading to lower Dice scores for active tumor and the tumor core. However, in several other patients rCBV improved tumor delineation, in particular in small HGG lesions which were not easily detected by hNMF. Improved tissue differentiation within brain tumors by adding MRSI to cMRI

has been reported before [159] and was also shown to be statistically significant for each class in our study. Despite its low spatial resolution, several metabolites are among the most specific features to detect tumor and necrosis: elevated levels of Lip, Lac and Cho are only expected to be found within the tumor core. In several HGG patients, voxels outside the pathological region were wrongfully labelled as tumor or necrosis when omitting MRSI. In addition, small areas of non-enhancing tumor were sometimes not denoted as tumor by hNMF without MRSI. Statistical significance for adding DKI to the full MP-MRI dataset was shown for the tumor core and complete tumor region, but not for viable tumor. As was reported in [234] for MD, inclusion of the DKI parameters improved edema segmentation in several patients. It also resulted in better differentiation of necrosis from edema in some cases. When considering only cMRI data, statistical significance of the higher Dice scores for full MP-MRI has been shown for each tissue class. Compared to the reduced MP-MRI datasets, a larger decrease in mean Dice scores was found: 9% for active tumor, 14% for the tumor core and 15% for the complete tumor.

For the LGG patients, although the mean Dice score of the whole tumor region was highest for the full MP-MRI dataset, statistical significance was only shown against no-cMRI, with a decrease in mean Dice score by 12%. As LGGs represent an earlier stage of the disease, hallmarks such as angiogenesis, hypoxia, necrosis and increased cell proliferation are uncommon. The tumor tissue is mostly differentiable on cMRI and DWI, but with low specificity. Additionally, for no-cMRI the tumor tissue was not properly separated from CSF in some LGG patients, as these tissue types are mostly differentiable based on FLAIR. Since statistical significance could not be shown for the other reduced MRI datasets, it seems that additional MRI modalities do not improve tumor differentiation in most LGG cases.

For both the HGG and the LGG patients, the mean correlation coefficient showed a less distinct decrease in most cases and even a slight increase in some cases compared to full MP-MRI. It must be noted that a higher correlation coefficient does not warrant better tissue differentiation: a significant amount of voxels will not contain purely one tissue type, such that the average feature vector per tissue region is expected to contain some contamination from other tissue types. Since the correlation coefficients are above 0.90 in most cases for the HGG as well as for the LGG patients, the pathological hNMF sources are representative for the corresponding tissue region and can be considered as patient- and tissue-specific patterns.

4.5.3 Tumor segmentation

Tumor segmentation is a challenging task in the image analysis of brain tumors. Manual delineation is still the gold standard in clinical practice. Nevertheless, it is a time-consuming and user-dependent process, prone to errors and with questionable reproducibility [165]. To overcome these limitations, many sophisticated algorithms have been developed for automated segmentation. However, few have achieved the speed, accuracy, and limited user interaction required for routine clinical use. One means of validation was to convert the hNMF results into an absolute voxel-by-voxel segmentation, by applying k-means clustering to the obtained abundance maps. Comparing the hNMF-based segmentation to the manual segmentation by an experienced radiologist, mean Dice scores for the whole tumor region of 83% and 85% were found for the HGG and the LGG patients, respectively. For the HGGs, mean Dice scores of 78% and 85% were found for the active tumor region and for the tumor core, respectively. These Dice scores are higher than the mean Dice scores of the best performing methods in [168] reporting on the BRATS challenge 2012-2013 results. It must be noted that these methods are mainly supervised classification methods which only rely on cMRI data. Nevertheless, it shows that our results are competitive with the state-of-the-art segmentation techniques, even when only cMRI data are considered. We have thus shown that the hNMF method presented in this study can be applied as a valid unsupervised segmentation technique. Another difference with the BRATS challenge is that we did not consider a full 3D imaging dataset, but only a 2D region corresponding to the MRSI ROI, as this is the only region where a full set of MP-MRI features was available. This is an important limitation of including MRSI, as tissue characterization of the entire tumor lesion is mandatory for treatment planning and response monitoring. Although 3D MRSI is available, it is often not possible to cover the entire tumor volume, as the MRSI spatial grid has to be positioned entirely within the brain to avoid artefacts like lipid contamination from the skull. Recently, an echo-planar MRSI sequence has been presented that does allow full coverage of the whole brain [164]. Full tumor coverage with MRSI will be possible when such sequences become available for clinical practice.

The relative contribution of each tissue type within each voxel is presented by means of the abundance maps. These maps allow direct visual interpretation of the hNMF results, combining the information of all MRI parameters into one map per tissue type. The relative contributions of different tissue types within the tumor region provide us with more insight in the spatial heterogeneity of the tumor. To further synthesize visual assessment, the abundance maps can be combined into a single nosologic image, in which each tissue class is represented by one colour and colour mixtures are used to quantify the relative contribution of each tissue type in each voxel [135, 154].

4.5.4 Unsupervised vs supervised classification

Given the wide range of tumor characteristics that is reported for HGGs and the large heterogeneity that is often seen within the same lesion, unsupervised methods have the advantage of not making any prior assumptions about the expected values of the individual features. This has allowed us to apply hNMF on a patient-by-patient basis, resulting in patient-specific tissue patterns. Several studies have applied supervised classification methods to characterize brain tumor substructures. Di Costanzo *et al.* applied linear discriminant analysis to MP-MRI data to perform ROI based classification of solid tumor, edema, edema/tumor and infiltrated peri-tumoral tissue on low- and high-grade gliomas [64]. Classification accuracy of 67% was found for solid tumor ROIs. Another study used support vector machines on a combination of cMRI and DTI to distinguish enhancing tumor, non-enhancing tumor and edema in high-grade brain tumors [259]. Voxel-wise classification rates of 89% and 34% were found for enhancing and non-enhancing tumor, respectively. As was pointed out in the latter study, tumor regions might contain varying degrees of neoplasm combined with healthy tissue or edema. Identifying tissue-specific patterns across patients is hampered due to the co-existence of different tissue types, especially when mixed tissue classes are considered. This aspect of tissue complexity is not taken into consideration by most other classification methods. One limitation of unsupervised methods such as (h)NMF, is that the sources are not automatically labelled. This was done based on visual interpretation of the sources and their abundance maps.

4.5.5 Voxel-wise vs ROI analysis

Previous advanced MRI studies have often considered single statistical values for a lesion, or for a few ROIs within the lesion. Given the heterogeneous nature of HGGs, a simple statistical metric such as the mean value is inadequate in characterizing the whole tumor. Numerous studies have proposed minimum or maximum values of PWI, DWI or MRSI parameters to assess tumor malignancy [40, 293] or to detect the most malignant part of the lesion for biopsy targeting [264, 268]. Such methods, although they have shown statistical significance in most cases, lack robustness against outlier values and have limited interpretability. A more detailed intra-tumoral tissue characterization will allow us to move towards more personalized and dedicated treatment planning.

4.5.6 Conclusion

In this chapter, a hierarchical NMF method was presented to characterize brain tumor heterogeneity based on MP-MRI data. hNMF basically consists of 2 levels of NMF. An iterative voxel selection procedure was used at the second level of NMF, to minimize cross-contamination between different tissue sources. To select the optimal result, a sparsity criterion was introduced which assumes that most voxels will contain predominantly one tissue type. hNMF provides tissue-specific signatures of the pathological tissue classes, as well as abundance maps of their spatial distribution. The abundance maps were converted into absolute tissue segmentation and good spatial alignment was found with manual segmentation by a neuro-radiologist. To assess the added value of the individual MRI modalities, hNMF analyses were also performed on reduced MRI datasets. For the HGG patients, Dice scores were significantly higher for most tissue classes with the full MP-MRI dataset compared to the reduced datasets. Dice scores were also highest with full MP-MRI for the LGG patients, but statistical significance could not be shown against most reduced MRI datasets. This suggests that inclusion of additional MRI modalities is mainly beneficial for HGG segmentation, as HGGs present with a higher degree of intra-tumoral heterogeneity and more pathological subcompartments.

Chapter 5

The successive projection algorithm as an initialization method for brain tumor segmentation using NMF

NMF has become a widely used tool for additive parts-based analysis in a wide range of applications. As NMF is a non-convex problem, the quality of the solution will depend on the initialization of the factor matrices. In this chapter, the successive projection algorithm (SPA) is proposed as an initialization method for NMF. SPA builds on convex geometry and allocates endmembers based on successive orthogonal subspace projections of the input data. SPA is a fast and reproducible method, and it aligns well with the assumptions made in near-separable NMF analyses. SPA was applied to the UZ Gent MP-MRI dataset for brain tumor segmentation using different NMF algorithms. Comparison with common initialization methods shows that SPA achieves similar segmentation quality and it is competitive in terms of convergence rate. Whereas SPA was previously applied as a direct endmember extraction tool, improved segmentation results have been shown when using SPA as an initialization method, as it allows further enhancement of the sources during the NMF iterative procedure. This study has been submitted as a journal paper for peer review as [220].

5.1 Introduction

Non-negative matrix factorization (NMF) has become a widely used tool for multivariate data analysis in the fields of blind source separation and pattern recognition. NMF decomposes a non-negative input matrix X into the product of 2 non-negative factor matrices W and H, providing a low-rank (rank r) approximation. As such, it reveals the basic components which are present in the data, the so-called sources, and models each input signal (i.e. column of the input matrix) as a weighted sum of the sources. The columns of W represent the sources and each column of H contains the weights, or so-called abundances, of the sources for one particular column of X. The most commonly used similarity measure for X and its factorization is the Frobenius norm. It will also serve as the cost function for solving the optimization problem:

$$\min_{W,H} f(W,H) = \min_{W,H} \quad \frac{1}{2}\|X - WH\|_F^2 \quad \text{, such that} \quad W \geq 0, \quad H \geq 0 \quad (5.1)$$

As NMF is a non-convex optimization problem, the obtained factorization will be a local rather than the global minimum of the cost function. The final result therefore depends on the initialization of the factor matrices, W_0 and H_0. Random initialization is the benchmark and is used in the vast majority of NMF studies. However, the quality and reproducibility of the NMF result is rarely questioned when using random initialization. As was pointed out in [19], one would have to run NMF with random initialization a sufficient number of times, then select the 'best' run based on some criterion (e.g. lowest Frobenius residual error), to warrant robust and reproducible NMF results. This is also how we proceeded in Chapter 4 for the initialization of the hNMF algorithm: hNMF analysis was repeated 20 times with random initialization for W_0 and H_0, then selected the optimal solution based on the H_{ratio3} criterion. Repetitive random initialization of course increases computation time dramatically.

Alternatively, more advanced initialization strategies have been suggested for NMF. Besides random initialization, other randomization based methods have been suggested. These methods obtain the columns of W_0 by averaging a number of random columns of X or a subset of X [132]. Despite having a low computational cost and providing a more realistic first estimate of the sources compared to random initialization, these methods also suffer from a lack of reproducibility. Another group of initialization schemes is based on low-rank approximation methods of the input data matrix, such as the singular value decomposition [19, 205, 289] or independent component analysis [285]. These methods rely on the most significant low-rank components and their corresponding source vectors to initialize W_0 and H_0. As these methods impose constraints such as orthogonality or statistical independence to the source

vectors, the non-negative structure of the input data is lost, introducing negative values into W_0 and H_0. Some straightforward ways of dealing with these negative values have been proposed, such as setting them to zero [19], replacing them by a mean value from the input matrix [19] or taking their absolute values [205]. Clustering algorithms have also been suggested for NMF initialization, such as spherical k-means clustering [275], fuzzy C-means clustering [211, 290] and subtractive clustering [38]. Clustering-based initialization schemes will provide more realistic source estimates compared to low-rank approximation methods, but they can become computationally expensive. Furthermore, clustering methods usually require some initialization themselves. Most of the proposed initialization methods have been compared with random initialization in terms of convergence rate and/or quality of the solution. However, different random initializations will lead to different NMF results, making it a questionnable reference. It is unclear how previous studies have coped with this lack of reproducibility.

NMF has been used in several biomedical applications, such as ECG-EMG signal unmixing [176], prostate tumor detection [134] and brain tumor segmentation [181, 225]. Brain tumor segmentation is a crucial task for planning surgical resection, for radiotherapy planning and to monitor tumor growth or shrinkage during follow-up [14]. It aims at outlining the total tumor volume as well as its main constituting tissue compartments, i.e. active tumor, necrosis and edema. NMF has been applied for tumor segmentation of MRI data [181, 225]. In this context, the columns of the input matrix X correspond to the different voxels and the rows represent the MRI parameters. The sources in W are interpreted as tissue-specific signatures and the abundances in H as the proportions of the different tissue types in each voxel. As such, NMF models each voxel's MRI feature set as a weighted sum of the tissue-specific signatures. The abundances associated with each source can be visualised as a segmentation of the image in various tissue components and evaluated by comparison against manual expert segmentation.

The Successive Projection Algorithm (SPA) is a forward selection method which minimizes collinearity of the selected variables in vector space. It was introduced by Araújo *et al.* [2] and has been commonly used for feature selection. Several studies have considered SPA as an endmember extraction tool for hyperspectral unmixing [80, 287]. SPA is fast and reproducible, and it takes advantage of the geometrical convexity that is seen in a wide range of NMF problems [80]. The current chapter proposes to use SPA as an initialization method for NMF. We illustrate the use of SPA initialization on an MP-MRI dataset, applying NMF for brain tumor segmentation. The performance of SPA in terms of segmentation quality and convergence rate is compared with other common initialization methods, i.e. random initialization, non-negative double

singular value decomposition (NNDSVD) and fuzzy c-means clustering (FCM). Non-deterministic methods such as random initialization are run repetitively, making sure the reported results are reproducible. Several NMF algorithms are discussed and their sensitivity to the initialization methods is investigated.

5.2 MP-MRI dataset

21 patients who were diagnosed with a high-grade glioma (HGG) underwent a multi-parametric MRI acquisition protocol at UZ Gent. The local ethics committee allowed a retrospective analysis of the data. The acquisition protocol consisted of cMRI, PWI, DWI and MRSI. As was shown in Chapter 4, combining these MRI modalities leads to improved brain tumor segmentation [225]. The MR examinations were performed on a 3T Siemens Trio Tim scanner (Erlangen, Germany), using a standard 12-channel phased array head coil. Table 5.1 gives an overview of the acquisition protocol and the obtained MP-MRI parameters for each MRI sequence. A detailed description of the acquisition protocol and data pre-processing methods for the UZ Gent dataset can be found in Section 3.2.

	UZ Gent dataset
cMRI	**T1, T1c, FLAIR**
PWI	*Gradient echo DSC-MRI* **rCBV**
DWI	*Spin echo DWI (b_0, b_{500}, b_{1000} in resp. 1, 3, 3 directions)* **MD, b0**
MRSI	*3D long TE ^{1}H-MRSI* **Lac, Lip1, NAA, Glx, Cre, Cho**

Table 5.1: Schematic overview of the MP-MRI protocols and derived parameters of the UZ Gent dataset.

To be able to perform voxel-wise analyses, all MP-MRI maps were co-registered and brought to the same spatial resolution. The T1c images were resliced such that the slices were parallel with the MRSI slabs, at a resolution of $1 \times 1 \times 3\text{mm}^3$. All other MP-MRI maps except the metabolites were rigidly co-registered to the resliced T1c reference using SPM8 (Wellcome Department of Imaging Neuroscience, University College London, UK). Skull-stripping was applied to all modalities prior to co-registration. The normalized mutual information criterion was used for co-registration [161], with cubic spline interpolation for reslicing. Co-registration was not possible for the MRSI data, due to their

low spatial resolution and the limited region of interest. The metabolite maps were spatially aligned with the T1c reference, assuming no patient motion had occurred between MRSI and T1c acquisition. Only voxels within the MRSI volume of interest (VOI) were included in the analyses, as only those voxels had a full MP-MRI feature set. The VOI consisted of 2 adjacent MRSI slabs which were manually positioned to include tumor, peri-lesional edema and normal brain tissue. The VOI extended over 10 consecutive image slices. One or two voxel bands at the outer edges of the spectroscopic grid were omitted from the analyses when showing severe chemical shift displacement effects or lipid contamination.

Twelve MP-MRI features were obtained from the raw acquired data after pre-processing: T1, T1c, FLAIR, rCBV, MD, b0, Lip1, Lac, NAA, Glx, Cre and Cho. In addition, 3×3 and 5×5 moving average windows were applied to the non-metabolic feature maps, and these averaged feature maps were also added to the feature set to improve robustness of the segmentation results [225]. A total of 24 MP-MRI features was finally obtained for each voxel.

5.3 NMF and initialization methods

5.3.1 NMF methods

aHALS NMF

Hierarchical alternating least squares (HALS) NMF is an exact block-coordinate descent method where blocks of variables (i.e. columns of W and rows of H) are optimized in a cyclic way [50]. HALS relies on the observation that finding the optimal factor matrix H when W is fixed, and finding the optimal W when H is fixed, are convex problems, as opposed to the original non-convex NMF problem as defined in Equation (5.1). See Section 2.1.2 for a detailed discussion of HALS NMF and its implementation. Considering the HALS algorithm, the most expensive computational step for one full iteration of updating W involves the computation of XH^T and HH^T. Gillis *et al.* suggested to perform multiple rounds of updating W before updating H, such that the expensive computations above don't need to be repeated at each iteration. The same reasoning can be made for updating H. This concept has been implemented in accelerated HALS (aHALS) NMF [82]. aHALS NMF is considered in this chapter, as it has improved computational efficiency while at least maintaining performance compared to the original HALS algorithm.

Convex NMF

Convex NMF [66] imposes the constraint that the source vectors (i.e. the columns of W) must lie within the column space of X. As such, each source is a weighted sum of the data points. An auxiliary matrix A is introduced to define these weights:

$$W = XA \quad \text{such that} \quad X \approx XAH \tag{5.2}$$

Multiplicative update rules are defined for A and H (see Section 2.1.2). Convex NMF comes with the appealing property that without any additional constraints, A and H are both sparse.

SDF NMF

Unlike most NMF implementations, such as HALS and Convex NMF, the structured data fusion SDF framework, as implemented in Tensorlab [260], does not solve the NMF problem by iteratively updating W and H through convex sub-formulations of Equation (5.1). Instead W and H are simultaneously updated using optimization theory. A transformation of variables is used to convert the constrained optimization problem in Equation (5.1) into an unconstrained problem, by squaring the entries of the factor matrices. The resulting non-linear least-squares problem is solved using the Gauss-Newton method with a dogleg trust region (GNDL) [177]. Gauss-Newton linearizes the quadratic problem using a Taylor expansion. It avoids expensive second-order computations for calculating the Hessian matrix. At each iteration of the GNDL algorithm, a step is calculated by iteratively solving the linearized version of Equation (5.1) using conjugate gradients.

hNMF

Hierarchical NMF (hNMF) has been introduced in the previous chapter for brain tumor characterization [225]. Similar hierarchical variants of NMF have also been considered for differentiating tumors in the prostate [134] and in the brain [153]. hNMF consists of a 2-level approach, and aHALS NMF is implemented at both stages. Tissue types which are most similar are assigned to the same source after a first level of rank-2 NMF. Tissue-specific sources are obtained after the second level of NMF. The sources are then recombined in a final step to calculate the tissue abundances using non-negative least squares fitting. A full description of the hNMF procedure can be found in Section 4.3. As hNMF consists of 2 levels of NMF, the initialization methods are applied at both NMF stages.

5.3.2 Geometrical interpretation of NMF

Before discussing the initialization methods, it is worthwhile to consider the NMF problem from a geometrical perspective. Let's consider a non-negative input matrix X of rank 2, containing n data points in three-dimensional space (i.e. $X \in R_+^{3\times n}$). The fact that X has rank 2 implies that all its data points belong to a two-dimensional subspace, and they are all located in the non-negative orthant (see Figure 5.1). To solve the NMF problem, we need to find 2 sources which will be located in the same two-dimensional subspace as the data points. Let's further assume that for each source, there is at least one data point in X that purely contains that source (i.e. the pure-pixel assumption). In that case, it can be seen that all data points are confined within a convex cone, whose edges intersect with the pure data points. Any data point within this cone can be obtained as a weighted sum of the extreme vectors of the cone. From this geometrical point of view, NMF comes down to finding the vertices of a convex cone spanning all the data points.

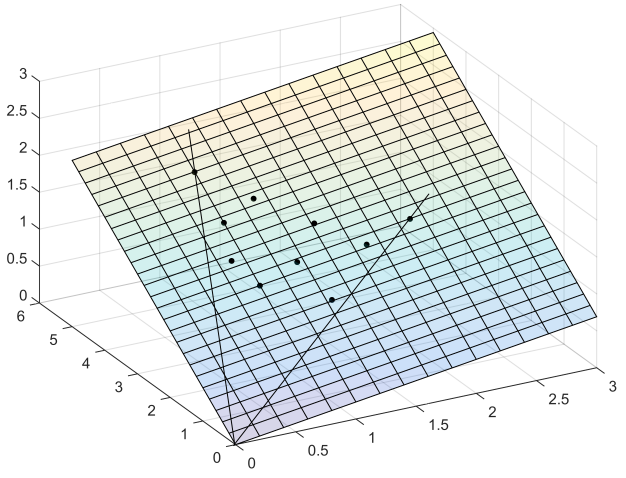

Figure 5.1: Geometrical representation of exact NMF for a rank-2 non-negative input matrix $X \in R_+^{3\times 10}$.

5.3.3 Initialization methods

Successive projection algorithm

SPA is a forward selection method which minimizes collinearity of the selected variables in vector space. SPA works as follows: out of a set of data points (i.e. the columns of X), SPA will first select the data point p_1 with the highest l_2-norm. All data points are then projected onto the orthogonal complement of p_1. The data point p_2 with the highest l_2-norm in this projected subspace has minimal collinearity to the previously selected point p_1. Therefore, p_2 is also selected and all data points are now projected onto the orthogonal complement of the subspace formed by p_1 and p_2. The next selected data point p_3 is found as the data point with the highest l_2-norm in the orthogonal subspace of p_1 and p_2, and so on. Algorithm 1 describes the SPA procedure.

Algorithm 1: SPA

Input: $X \in R_+^{m \times n}$, rank r
Output: Set of vectors $\{p_1, p_2, ..., p_r\} \in X$

1. Initialize matrix $S = X$
2. **for** $j = 1 : r$
3. find index i^* such that $i^* = \text{argmax}_{i=1,...,n} \parallel S(:, i) \parallel_2^2$
4. set $p_j = X(:, i^*)$, $s_j = S(:, i^*)$
5. update $S \leftarrow (I - \frac{s_j s_j^T}{\|s_j\|_2^2})S$
6. **end**

When the set of data points is bounded by a convex hull, SPA will return the vertices of the convex hull as an output. As we have seen in the previous section, the data points of a non-negative input matrix X are bounded by a convex cone, and under the pure-pixel assumption the vertices of this cone correspond to the NMF sources. SPA therefore aligns well with the objective of NMF, and has been commonly considered as an endmember extraction tool for NMF [80, 171]. In this chapter, SPA will be considered as an initialization method. The columns of W_0 will be formed by p_1, p_2,..., and p_r. SPA does not directly provide an initialization for the abundance matrix, H_0. Based on W_0, non-negative least squares fitting was used to obtain H_0. For Convex NMF, an initialization of the auxiliary matrix A_0 is required. Solving Equation (5.2) for

A, we find:

$$A \approx H^T(HH^T)^{-1} \tag{5.3}$$

Since A must be non-negative, we obtain A_0 as proposed in [66]:

$$A_0 = A^+ + 0.2E\langle A^+ \rangle \tag{5.4}$$

with A^+ being the positive part of A, $\langle A \rangle = \sum_{n,k} |A_{n,k}|/\|A_{n,k}\|_0$ and $\|A_{n,k}\|_0$ is the number of nonzero elements in A. E is a matrix of ones with the same dimensions as A. For the other initialization methods discussed below, A_0 is obtained from H_0 in a similar way.

Repetitive random initialization (rRandom)

Initialization obtained from SPA is compared with rRandom initialization, where the elements of W_0 and H_0 are set to uniformly distributed non-negative values between 0 and 1 at each run. We ran NMF 30 times with random initialization to obtain reproducible results. For rRandom, the final NMF result is selected as the random run with the lowest residual error as defined by the objective function in Equation (5.1).

Non-negative double singular value decomposition (NNDSVD)

NNDSVD is based on 2 levels of SVD. First, the best rank-r approximation of the input data matrix X is computed based on truncated SVD:

$$X \approx \sum_{i=1}^{r} \sigma_i C_i \quad \text{such that} \quad C_i = u_i v_i^T \tag{5.5}$$

with u_i and v_i being the ith left and right singular vectors of X. As we require a fully non-negative initialization, the positive component of C_i, C_i^+ is withheld, being the nearest non-negative rank-2 approximation of C_i. It was shown that the dominant singular triplet of C_i^+ can easily be obtained by decomposing the singular vectors u_i and v_i into their positive and negative components. W_0 and H_0 are then initialized based on the dominant singular triplet of each C_i^+. A detailed description of the NNDSVD algorithm can be found in [19]. Several studies have considered NNDSVD for NMF initialization [38, 101]. It must be noted that NNDSVD in its original form introduces a significant number of zero elements in W_0 and H_0. This can be problematic for some NMF algorithms, i.e. elements initialized at zero will remain zero in the final solution. This is

for instance the case for the original multiplicative update NMF algorithm proposed by Lee and Seung [148], and also for the SDF NMF algorithm that we are considering. To overcome this kind of convergence problem, we set all zero elements in W_0 and H_0 to a small positive value, i.e. to 1% of the mean value of the positive elements of W_0 and H_0, respectively.

Fuzzy c-means clustering (FCM)

FCM aims at partitioning the input data into c fuzzy clusters. The data points have a degree of belonging to each of the clusters centers, as in fuzzy logic, rather than belonging completely to just one cluster. The degree of membership of a data point i belonging to cluster j is defined by a weighting value w_{ij}, which depends on a distance metric between the data point and the cluster centroid (e.g. Euclidean distance). FCM is an iterative algorithm, in which the cluster centroids and the weighting values are iteratively updated until convergence [179]. FCM has been applied as an initialization method for NMF in several studies [211, 290], with the cluster centroids being assigned to the columns of W_0 and the membership weights per cluster serving as the rows of H_0. It has to be noted that FCM in itself is non-deterministic, meaning that the final output will depend on the initialization. Therefore, we applied the same approach as with random initialization: run NMF with randomly initialized FCM 30 times and select the run with the lowest residual Frobenius norm as defined in Equation (5.1) as the final solution.

5.3.4 Validation

Segmentation of the pathological tissue types, i.e. actively growing tumor, necrosis and edema, was obtained from NMF by applying k-means clustering to the abundance values of H. As it is assumed that each source and its associated abundances correspond to one tissue type, we initialized the cluster centroids by setting one abundance value to 1 and all others to 0. The NMF segmentations are compared with manual segmentation performed by an experienced neuroradiologist using MRIcron [213]. The overlap between segmentations was quantified using the Dice score (see Section 2.2.1). Dice scores were calculated for 3 tissue classes, as defined in the BRATS challenge [168]: active tumor, the tumor core (i.e. active tumor and necrosis) and the whole tumor (tumor core and edema). To compare performance of using SPA as an endmember extraction tool to using it as an initialization method, segmentation results were also assessed for direct SPA output, i.e. based on W_0 and H_0. We will denote the NMF results obtained with the different initialization methods as NMF_{SPA}, NMF_{NNDSVD}, NMF_{FCM} and NMF_{rRandom}, respectively.

5.4 Results

Figure 5.2 shows an example of a co-registered image set of a GBM patient. The segmentation results for active tumor and necrosis obtained with the 4 different NMF methods using SPA are also shown (in blue), and are compared with the manual segmentation by the radiologist (in green). The single-level NMF methods show some active tumor segmentation outside of the pathological region, which is less the case for hNMF. All NMF methods obtained good segmentation correspondance for necrosis, except that aHALS NMF showed some underestimation of the necrotic area.

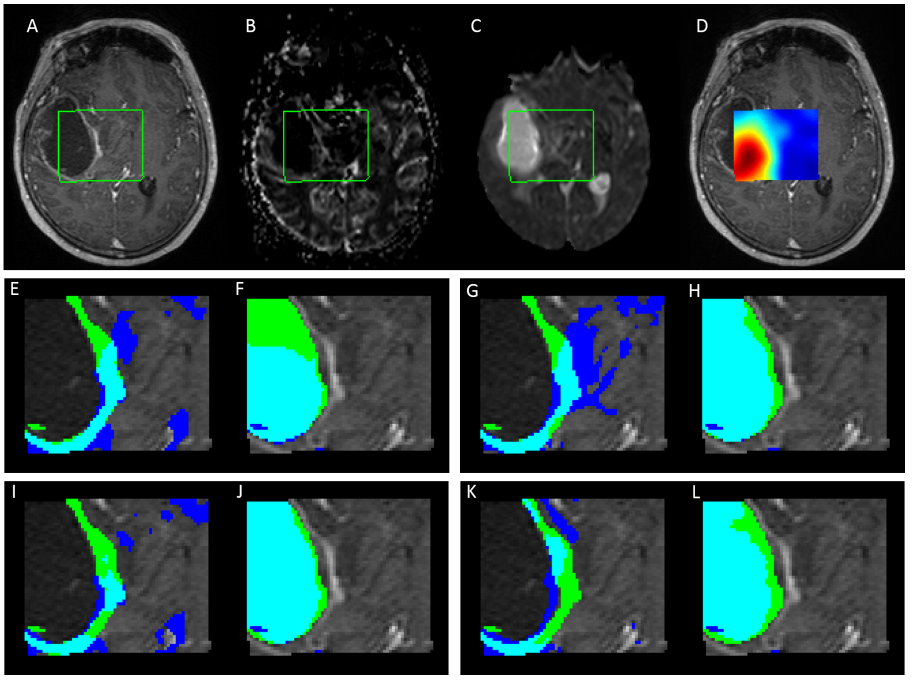

Figure 5.2: Co-registered set of MRI images, showing T1c(A), rCBV(B), MD(C) and Lac(D). The green frame in the images of the upper row marks the MRSI region of interest. Segmentation results are shown for active tumor and necrosis, respectively, for aHALS NMF(E,F), Convex NMF(G,H), SDF NMF(I,J) and hNMF(K,L) with SPA initialization. NMF segmentation is shown in blue, manual segmentation by a radiologist in green and segmentation overlap in cyan.

Table 5.2 shows the mean Dice scores with their standard deviation per tissue

class for SPA, NNDSVD, FCM and rRandom initialization. Looking at the results per tissue class and per NMF method, similar mean Dice scores are found with the different initialization methods, with variations mostly below 5%. The largest difference is found for active tumor with Convex NMF and SDF NMF, where NMF_{FCM} gives 5% lower Dice scores than NMF_{SPA}. Mean Dice scores for NMF_{NNDSVD} are lower than or at best equal to NMF_{SPA} for all tissue classes and for all NMF methods, but differences are never higher than 3%. Statistical significance of the differences could not be shown in any case using a two-tailed Wilcoxon signed rank test. Similar segmentation results are also found among the different NMF methods, although the highest Dice scores are consistently found for hNMF. The last column of Table 5.2 shows the segmentation results obtained directly from SPA output (W_0,H_0). The Dice scores from SPA were lower or at best equal to those from the NMF_{SPA} methods for each tissue class. Statistical significance of the higher Dice scores for NMF_{SPA} compared to direct SPA was found in about half of the cases.

		aHALS NMF	**Convex NMF**	**SDF NMF**	**hNMF**	**SPA (W_0,H_0)**
Dice active tumor [%]	**SPA**	65 ± 13	64 ± 18	65 ± 21	$69 \pm 15^*$	64 ± 20
	NNDSVD	63 ± 18	63 ± 19	62 ± 18	67 ± 17	-
	FCM	66 ± 14	59 ± 23	60 ± 21	69 ± 15	-
	rRandom	64 ± 19	63 ± 20	63 ± 19	70 ± 14	-
Dice tumor core [%]	**SPA**	76 ± 11	74 ± 14	$75 \pm 14^*$	$78 \pm 12^*$	73 ± 13
	NNDSVD	74 ± 12	74 ± 12	74 ± 11	76 ± 12	-
	FCM	76 ± 11	73 ± 11	72 ± 14	77 ± 13	-
	rRandom	76 ± 11	73 ± 15	75 ± 11	78 ± 12	-
Dice whole tumor [%]	**SPA**	78 ± 12	$84 \pm 10^*$	$80 \pm 11^*$	$86 \pm 8^*$	77 ± 13
	NNDSVD	78 ± 14	82 ± 10	80 ± 12	84 ± 9	-
	FCM	77 ± 14	82 ± 10	82 ± 12	84 ± 10	-
	rRandom	78 ± 14	82 ± 13	79 ± 13	85 ± 8	-

Table 5.2: Comparison of the mean Dice scores and their standard deviation between different initialization methods.* indicates statistically significantly higher Dice scores with SPA initialization compared to direct SPA endmember extraction (right column), using a one-tailed Wilcoxon signed rank test ($p < 0.05$).

Table 5.3 shows the mean number of iterations to convergence for the 3 single level NMF methods using the different initialization schemes. These values could not directly be reported for hNMF, as it consists of 2 levels of aHALS NMF. The same stopping criteria were applied to all NMF methods: a maximum number of iterations of 10000 and a convergence tolerance of 10^{-5} for the relative difference of two subsequent values of the residual error $\|X - WH\|_F$. For NMF_{rRandom} and NMF_{FCM} , the number of iterations is reported for the selected run, i.e. the run with the lowest final residual error. aHALS NMF has fast convergence,

with the mean number of iterations below 200 for all initialization methods. Convergence rate is low for Convex NMF, where all initialization methods result in a mean number of iterations close to 9000. Only for SDF NMF, we see a clear difference in convergence rate among the initialization methods. NMF_{SPA} and NMF_{NNDSVD} have the fastest convergence, with a mean number of iterations of 218 and 175, respectively. NMF_{rRandom} requires on average 418 iterations to reach convergence, which is about twice as many. NMF_{FCM} has the lowest convergence rate, with a mean number of iterations of 832. In a previous study comparing different initialization methods for document mining, NMF with FCM and random initialization were also found to converge relatively slowly compared with NNDSVD [38].

	#Iterations			
	SPA	**NNDSVD**	**FCM**	**rRandom**
aHALS NMF	179	181	140	190
Convex NMF	8923	8796	8819	8810
SDF NMF	218	175	832	418

Table 5.3: Mean number of iterations to reach convergence for the different initialization methods. Convergence tolerance was set to 10^{-5} and the maximum number of iterations to 10000.

Figure 5.3 illustrates the convergence behavior of the 3 single-level NMF methods when comparing initialization schemes for a particular GBM patient. Initial errors are lowest for SPA and FCM. The main reduction of the residual error occurs within the first 100 iterations in all cases, except for Convex NMF with rRandom. For aHALS and Convex NMF, all initialization methods converge to approximately the same residual error, whereas for SDF NMF, SPA and FCM converge to a higher final residual.

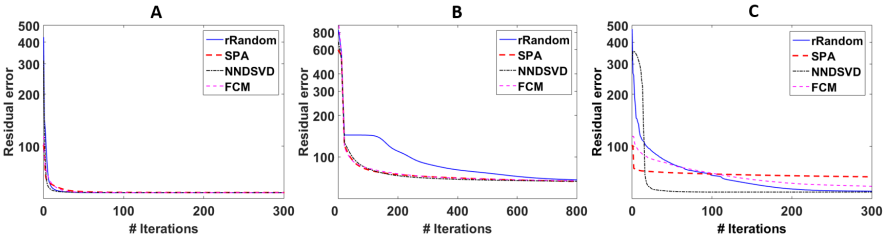

Figure 5.3: Convergence plots for aHALS NMF (A), Convex NMF (B) and SDF NMF (C) with the different initialization methods. The residual error, $\|X - WH\|_F$, is shown on a log scale. For rRandom and FCM, the shown curve corresponds to the selected run with the lowest residual error.

5.5 Discussion

5.5.1 NMF performance

This study illustrates the behaviour of SPA as an initialization method for NMF in comparison to other common initialization methods for brain tumor segmentation. As noted in [19], there is no consensus on how to assess NMF performance when comparing different initialization methods. Many studies report the final residual error as a measure of performance [38, 132, 205], which has the advantage of being a generally applicable measure. Other studies use a validation which is specific to the problem at hand. Ortega-Martorell *et al.* calculated correlation coefficients as a quality measure of the obtained tissue sources for brain tumor characterization using single-voxel magnetic resonance spectroscopic data [181]. Wild *et al.*, besides looking at the residual error, assessed the quality of image sources for a facial image reconstruction problem based on visual inspection [275]. We assessed quality of the segmentation results using the Dice score, because a lower residual error does not warrant a better parts-based solution. For instance, one can observe in Figure 5.3 that higher residual errors remain for Convex NMF compared to aHALS and SDF NMF. This is due to the additional constraint imposed by Convex NMF that the sources have to be a weighted sum of the input data. As can be seen in Table 5.2, Convex NMF reaches similar Dice scores as the other single-level NMF methods. Therefore we decided to assess the quality of the NMF result using direct validation of the problem at hand, i.e. segmentation of the pathological tumor regions. For all considered NMF methods, mean Dice scores were similar among the initialization methods and differences were not higher than 5%. Statistical significance of the differences has not been found, showing that SPA achieves similar performance compared to some of the most common initialization methods.

For most advanced initialization methods, similar performance has been reported compared to random initialization, although residual errors generally tend to be rather higher than lower. Wild *et al.* discussed how most advanced methods obtain a head start in terms of residual error compared to random initialization, but this advantage is mostly lost towards the end of the iteration process [275]. In some cases, random initialization will achieve a lower error towards the end of the convergence process, which is also illustrated in [275] and in Figure 5.3 for SDF NMF. This is explained by the fact that advanced initialization often pushes the NMF result in a certain direction, thereby imposing restrictions on the factorization.

5.5.2 Convergence and computational cost

Although a head start in terms of the initial error has been reported for most advanced initialization methods compared to random initialization, mixed results have been found in terms of the number of iterations to reach convergence [38]. As shown in Table 5.3, for aHALS NMF and Convex NMF we did not find considerable differences in convergence rate among the different initialization methods. Only for SDF NMF, SPA and NNDSVD converge significantly faster than rRandom and FCM. Besides convergence speed, another important concern regarding computation time is the fact that random and FCM initialization are non-deterministic. Therefore, individual runs of these initialization strategies will not provide reproducible results, and some particular initializations might lead to a local optimum which is far from the global optimum. We decided to run NMF with repetitive random and FCM initialization, with a sufficient number of runs such that results were reproducible. This of course dramatically increased computation times for NMF_{rRandom} and NMF_{FCM}.

Concerning computational complexity, SPA is a straightforward and computationally inexpensive method, requiring a total of $6mnr$ operations [80]. For NNDSVD, the main computational cost comes from calculating the truncated SVD of the input data matrix X. The SVD is commonly calculated by means of some iterative algorithm, with a computational complexity of $O(mnr)$ operations per iteration [19]. This means that the leading constant, which depends on the number of iterations, will usually be much higher for NNDSVD compared to SPA. Similarly, FCM is an iterative procedure with a complexity of $O(mnr)$ per iteration [38], such that the leading constant will be higher than for SPA. It must be noted, however, that both NNDSVD and FCM deliver W_0 and H_0 initialization, whereas SPA only delivers W_0. As other initialization methods also only provide W_0, several suggestions have been made on how to additionally initialize H_0. Random initialization is sometimes used for H_0 in combination with a more advanced initialization of W_0 [132], although this partially annihilates the advantages of using a fixed and input-specific initialization. Another option, which we applied in the current study, is to use non-negative least-squares fitting to initialize H_0 based on W_0 and the input matrix X. We used the standard Matlab routine lsqnonneg, which is based on the relatively slow algorithm as originally described in [141]. Improved algorithms have been proposed in the meantime [27]. Another possibility is to initialize H_0 based on least-squares fitting without non-negativity constraints, then setting all negative values to zero or a small positive value. This actually comes down to one updating step of the basic block coordinate descent ALS algorithm. Standard least-squares will produce H_0 with a higher initial error, but at a significantly lower computational cost.

5.5.3 SPA: initialization or direct source extraction

The current study is not the first one to combine SPA with NMF. Several studies have applied SPA as an endmember extraction algorithm for NMF in hyperspectral unmixing [80, 171]. However, to the authors' knowledge, SPA has so far not been considered as an initialization strategy. We hypothesized that there might be some benefit to assigning the SPA sources to W_0 instead of directly using them as the final columns of W, because SPA will not always directly provide the most suitable sources for the NMF problem at hand. One of the main limitations of SPA is its sensitivity to outliers. As SPA looks for the outer vertices of a geometrical hull, outliers might be wrongly assigned as endmembers. Several methods have been proposed to either make SPA less sensitive to outliers [287] or to discard the outliers beforehand [84]. Another important caveat is that SPA assumes the NMF problem to be near-separable [80], which is equivalent to the spatial representation of the input data by a convex hull. The assumptions that are intrinsic to near-separable NMF, i.e. the pure-pixel assumption and the sum-to-one constraint on the abundances, will often not hold, especially in the case of highly mixed NMF problems. But when applying SPA as an initialization method, the NMF procedure might still correct the source estimates. For the MP-MRI dataset considered in this study, the pure-pixel assumption holds for most of the MRI modalities due to their high spatial resolution. However, the poor resolution of the MRSI data leads to significant partial volume effects, such that the pure-pixel assumption does not hold for these data. We can further assume that the sum-to-one constraint approximately holds, as each voxel represents a fixed tissue volume. By comparing the Dice scores of NMF_{SPA} to those obtained directly from SPA output matrices W_0 and H_0, it was shown that we can obtain better segmentation results when using SPA as an initialization method rather than as a direct source extraction tool. Mean SPA Dice scores were lower than the Dice scores obtained with any of the NMF methods and for all tissue classes. Statistical significance of the higher Dice scores for NMF_{SPA} could be shown for about half of the *(NMF method, tissue class)*-combinations in Table 5.2.

5.6 Conclusion

This study has proposed SPA as an initialization method for NMF. Whereas SPA has previously been applied as a direct endmember extraction algorithm, using it as an initialization method might allow further enhancement of the sources during the NMF iterative procedure. This advantage has been shown on an MP-MRI dataset for brain tumor segmentation. Compared to other initialization

methods, SPA is fast and provides a realistic and reproducible initialization. SPA shows similar performance in terms of segmentation quality of the pathological tissue classes, and it is competitive with the other initialization methods in terms of convergence rate. As the feasibility of using SPA initialization has been shown for this particular multi-parametric image segmentation problem, we encourage the exploration of SPA in other NMF applications as well.

Chapter 6

Comparison of unsupervised classification methods for brain tumor segmentation using MP-MRI data

Tumor segmentation is a particularly challenging task in high-grade gliomas (HGGs), as they are amongst the most heterogeneous tumors in oncology. As we have seen in Chapter 4, segmentation of HGGs and their subcompartments can benefit from the inclusion of advanced MRI modalities. In this chapter, we compare the performance of several unsupervised classification methods for HGG segmentation based on MP-MRI data including cMRI, DWI, MRSI and PWI. Compared to Chapter 4, the hierarchical NMF (hNMF) algorithm has been adapted to improve computational efficiency, and advanced initialization strategies are considered to further reduce computation times. All unsupervised methods are applied to the HGG patient cohorts of UZ Leuven and UZ Gent. This study has been published in [224].

6.1 Introduction

HGGs are amongst the most heterogeneous tumors in oncology. They are diffuse, exhibiting unclear and irregular boundaries, preferentially invading the surrounding tissue along white matter tracts [203]. Different stages of

the disease can occur within the same lesion [190], with varying degrees of tumor enhancement, mitotic activity and necrosis. Furthermore, tumor structures vary considerably in terms of size, shape and location. This makes tumor segmentation, i.e. the imaging-based delineation of the tumor and its subcompartments, particularly challenging in HGGs. Tumor segmentation is a crucial task in the treatment planning and follow-up of HGG patients. Manual segmentation by a neuro-radiologist, usually based on conventional MRI, is currently still the gold standard. It is a tedious and time-consuming task, and it is susceptible to subjective interpretation. Intra- and inter-observer variabilities of 20% and higher have been reported for manual tumor delineation [168, 271].

To alleviate these limitations and due to the high clinical relevance, there has been an increasing interest in the machine learning community for automated brain tumor segmentation. Classification methods, which rely primarily on voxel intensity-based features for differentiating tissue types, have been receiving the most attention [168]. They have the advantage of being able to deal with multi-parametric imaging datasets, combining voxel-wise information from different MRI modalities. Classification methods can be further categorized into supervised and unsupervised methods. Supervised classification methods naturally incorporate large amounts of prior knowledge in the form of a training dataset with known tissue labels. From this training dataset, the algorithm learns decision boundaries between tissue classes in high-dimensional feature space, which can then be applied to unlabeled test data. Supervised classification methods require extensive training datasets, to account for the wide heterogeneity in tumor appearance among glioma grades. Unsupervised classification methods reveal structure in a dataset by modeling the similarity within the data itself. These methods can directly be applied to any imaging dataset, irrespective of the acquisition protocol and without the need for any training data. Due to the lack of manually annotated ground truth, unsupervised methods depend more strongly on the incorporation of additional prior knowledge, e.g. by imposing spatial coherence [175] or feature-specific knowledge [119], to achieve a valid tissue segmentation.

Fuzzy c-means clustering (FCM) is one of the most popular unsupervised algorithms for brain tumor segmentation [87]. FCM takes into account the MRI feature overlap between tissue classes, assigning fuzzy membership values to different tissue types. FCM was introduced for brain tumor segmentation by Philips *et al.* [192] and was later combined with knowledge-based techniques for improved performance [73, 119]. Several studies have applied finite mixture modeling for unsupervised tumor segmentation. Mostly, each tissue type is modeled by a multi-variate Gaussian distribution, resulting in a Gaussian mixture model (GMM) [169, 288]. Recently, GMM has been shown to be competitive with state-of-the-art supervised classification alorithms for GBM

segmentation based on cMRI data [115]. Over the last years graph-cut based methods have become popular as well. A semi-supervised procedure based on graph-cut has been presented to differentiate tumor tissue from healthy brain using MRSI data [88]. Padole *et al.* have proposed a normalized-cut method using the mean-shift algorithm for increased computational efficiency [188].

The proposed segmentation algorithms have so far mainly been applied to cMRI data [73, 115, 169, 192, 288]. However, along with the recent emergence of advanced MRI modalities such as DWI, PWI and MRSI, numerous studies have suggested that tumor characterization might benefit from the additional structural, biological and biochemical information provided by these advanced MRI modalities [14, 39, 187]. Thus far, only few studies have combined different MRI modalities for glioma segmentation. Di Costanzo *et al.* used DWI, PWI and ^{1}H-MRSI to classify a set of manually delineated regions of interest (ROIs) in the non-necrotic part of the tumor and the peritumoral region [64]. Step-wise linear discriminant analysis was applied to the ROI averaged MP-MRI parameters to differentiate between regions of gross tumor, edema, tumor/edema, tumor infiltration and normal tissue. Highest classification accuracy was reported when all MRI modalities were included. Verma *et al.* applied support vector machines to a combined set of cMRI and DTI parameters to segment intensity-based tissue profiles on a voxel-wise basis [259]. They provide tissue probability maps as well as hard segmentation of enhancing tumor, non-enhancing tumor and edema in the pathological region. Zikic *et al.* applied a decision forest classifier on a set of voxel-wise intensities and context-aware features from cMRI and DTI [292]. They report an improvement of the segmentation accuracy when DTI is included compared to using only cMRI data. Kazerooni *et al.* are among the first to explore unsupervised classification methods for MP-MRI based glioma segmentation [119]. They applied spatial fuzzy c-means clustering to a combination of cMRI, DWI and PWI data, followed by a region growing step.

NMF methods have been previously proposed for differentiating brain tumors from healthy tissue using MRSI data [153, 181]. In this paper, we consider a hierarchical NMF variant (hNMF) and compare its performance to several state-of-the-art unsupervised classification algorithms for brain tumor segmentation using MP-MRI data. As extensive MP-MRI datasets might be more prone to variations in the acquisition protocol, unsupervised methods benefit from not requiring a uniform training dataset. hNMF has been introduced in Chapter 4, where HGG segmentation has been shown to improve significantly when using the MP-MRI data instead of using cMRI data only [225]. Compared to Chapter 4, the hNMF algorithm has been adapted to improve computational efficiency, making it approximately 10 times faster while maintaining segmentation performance. We will compare hNMF to 5 other unsupervised classification methods: 2 single-level NMF methods (aHALS NMF and Convex NMF) and 3

clustering methods (fuzzy c-means, Gaussian mixture modeling and spectral clustering). In order to make our conclusions more general, all methods are applied to the 2 independent MP-MRI datasets from UZ Gent and UZ Leuven.

6.2 MP-MRI datasets

The 2 MP-MRI datasets of UZ Gent and UZ Leuven are considered in this chapter. Both datasets included cMRI, DWI, PWI and ^{1}H-MRSI, but the acquisition protocols and the resulting sets of MRI features were different. Table 6.1 gives a concise overview of the acquisition protocols and MP-MRI features from both datasets.

	UZ Gent dataset (21 HGG patients)	**UZ Leuven dataset** (14 HGG patients)
cMRI	**T1, T1c, FLAIR**	**T2, T1c, FLAIR**
PWI	*DSC-MRI* **rCBV**	*DSC-MRI* **rCBV**
DWI	*DWI (3 b-values, 1+3+3 directions)* **MD, b0**	*DKI (4 b-values, 1+25+40+75 directions)* **MD, MK, FA**
MRSI	*3D long echo ^{1}H-MRSI* **Lac, Lip1, NAA, Glx, Cre, Cho**	*2D short echo ^{1}H-MRSI* **Lac, Lip1, NAA, Glx, Cre, Cho, Gly, mI**

Table 6.1: Schematic overview of the MP-MRI protocols and derived parameters of the 2 datasets acquired at UZ Gent and UZ Leuven.

6.2.1 UZ Gent

21 HGG patients (12 GBMs, 1 grade III astrocytoma, 1 grade III oligodendroglioma, 7 grade III oligo-astrocytomas) were included in the study. Retrospective analysis of the data was approved by the local ethics committee. MR examinations were performed on a 3T Siemens Trio Tim scanner (Erlangen, Germany), using a standard 12-channel phased array head coil. cMRI included T1 with and without contrast (T1c and T1, voxel size $0.9 \times 0.9 \times 0.9\text{mm}^3$) and FLAIR imaging (voxel size $1.0 \times 1.0 \times 1.0\text{mm}^3$). DWI consisted of a fast single-shot gradient-echo EPI sequence with diffusion gradient b-values of 0, 500 and 1000s/mm^2 acquired in 3 orthogonal directions (voxel size $2.0 \times 2.0 \times 3.0\text{mm}^3$). ADC maps were computed using weighted linear least squares fitting [258]. MD was obtained as the averaged ADC over the 3 directions. The T_2-weighted b0 images were also added to the feature set. PWI was performed by using a T_2^*-weighted dynamic susceptibility-weighted contrast-enhanced MRI (DSC-MRI) sequence (voxel size $1.8 \times 1.8 \times 5\text{mm}^3$). rCBV maps were quantified from the dynamic signal intensity curves using the method proposed by Boxerman *et*

al., which accounts for leakage correction [20]. A 3D ^{1}H-MRSI protocol with long TE was included (voxel size $10 \times 10 \times 15\text{mm}^3$). In the two-slice MRSI examination, a volume of interest was positioned manually to include tumor, perilesional edema and normal brain tissue. Metabolite quantification was performed using AQSES-MRSI [57], resulting in Lac, Lip1, NAA, Glx, Cre and Cho maps. A more detailed description of the UZ Gent MP-MRI acquisition protocol and pre-processing of the MRI features can be found in Chapter 3.

6.2.2 UZ Leuven

As we are considering segmentation of high-grade gliomas, only the 14 HGG patients of the UZ Leuven patient cohort were enrolled in the study (11 GBMs, 2 grade III astrocytomas and 1 grade III oligoastrocytoma). Written informed consent was obtained from every patient before participation. MRI acquisition was performed on a 3T Philips Achieva scanner (Best, The Netherlands), using a body coil for transmission and an 8-channel head coil for signal reception. cMRI consisted of T_2-weighted imaging (voxel size $0.45 \times 0.45 \times 4\text{mm}^3$), T_1-weighted imaging after contrast administration ($0.98 \times 0.98 \times 1\text{mm}^3$) and FLAIR ($0.45 \times 0.45 \times 4\text{mm}^3$). An extensive DWI EPI sequence was used to acquire DKI data (voxel size $2.5 \times 2.5 \times 2.5\text{mm}^3$). Diffusion and kurtosis tensors were estimated in each voxel using a constrained weighted linear least-squares algorithm [258]. MD, FA and MK maps were calculated according to [146, 196]. PWI consisted of a DSC-MRI sequence (voxel size $1.56 \times 1.56 \times 3\text{mm}^3$). rCBV maps were derived from the dynamic signal intensity curves using the method by Boxerman *et al.* [20]. Short TE ^{1}H-MRSI data were acquired for a 2-dimensional ROI positioned in the center of the tumor (voxel size $5 \times 5 \times 10\text{mm}^3$). AQSES-MRSI was used to quantify Lac, Lip1, NAA, Glx, Cre, Cho, Gly and mI. For a detailed description of the UZ Leuven acquisition protocol and pre-processing of the MRI features, the reader is referred to Chapter 3.

6.2.3 Image co-registration and voxel selection

To allow for voxel-wise segmentation, all MP-MRI parameters were co-registered and resampled to the same spatial resolution of $1 \times 1 \times 3\text{mm}^3$. cMRI data were skull-stripped and T1c served as a reference for rigid co-registration in SPM8 (Wellcome Trust Centre for Neuroimaging, University College London, UK), using the normalized mutual information criterion [161] and cubic B-spline interpolation for reslicing. The MRSI data were spatially aligned and resliced to the T1c reference. Only voxels within the MRSI ROI have a full set of MP-MRI features, so only those voxels were included in the segmentation analyses. For

the UZ Gent dataset, voxels from 10 consecutive slices located within the 3D MRSI ROI were included. For the UZ Leuven dataset, voxels in a central slice intersecting the 2D MRSI ROI were considered. Additional intensity-based features were added to the feature set to include localized spatial information. An in-plane local neighborhood of 3×3 and 5×5 voxels was used to calculate average intensity values that were assigned to the central voxel. These spatially averaged intensity values were added for all MP-MRI features except for the MRSI data, because of the low spatial MRSI resolution. Finally, each feature's full range was rescaled linearly to [0-1].

6.3 Methods

Several state-of-the-art unsupervised classification methods are applied and evaluated on the UZ Gent and UZ Leuven datasets for MP-MRI based brain tumor segmentation. Three NMF methods (aHALS NMF, Convex NMF and hNMF) and 3 clustering methods (fuzzy c-means, Gaussian mixture modeling and spectral clustering) are considered.

6.3.1 Non-negative matrix factorization

A detailed description of the NMF model can be found in Section 2.1. NMF factorizes a non-negative input matrix X into the product of 2 non-negative factor matrices W and H. In terms of MP-MRI based tissue segmentation, each column of X corresponds to one voxel's MP-MRI feature vector. The columns of W, the so-called sources, will represent tissue-specific MP-MRI signatures. Each row of H will then contain the relative abundances of one tissue type (i.e. the corresponding source of W) for all the voxels. The abundances in H tell us which tissue types are most prominent in each voxel. To convert the multi-linear NMF result into a hard segmentation, k-means clustering is applied to the abundance values in H. As we assume the NMF sources to correspond to the tissue classes, we initialize each cluster centroid with a different abundance value set to 1, and all other abundance values set to 0.

aHALS NMF

Hierarchical alternating least squares (HALS) NMF is an exact block-coordinate descent method where the columns of W and the rows of H are optimized in a cyclic way [50]. HALS relies on the observation that finding the optimal factor matrix H when W is fixed, and finding the optimal W when H is fixed, are

convex problems, as opposed to the original non-convex NMF problem. See Section 2.1.2 for a detailed discussion of HALS NMF and its implementation. Accelerated HALS (aHALS) NMF is a computationally more efficient variant of the original HALS NMF algorithm [82]. It iteratively performs several rounds of updating W before updating H, such that the most expensive computational steps for updating W (i.e. the computation of XH^T and HH^T) don't have to be repeated at each iteration. The same reasoning is applied for updating H.

Convex NMF

Convex NMF imposes the source vectors of W to lie within the column space of X [66]. Moreover, each source must be a weighted sum of some of the data points. An auxiliary non-negative matrix A is introduced to enforce this additional constraint:

$$W = XA \quad \text{such that} \quad X \approx XAH \tag{6.1}$$

Multiplicative update rules are defined to alternatingly revise A and H towards convergence (see Section 2.1.2). Convex NMF comes with the appealing property that without any additional constraints, A and H are both sparse.

Hierarchical NMF

Hierarchical NMF variants have been previously used for document clustering [128], for unmixing hyperspectral images [83] and for tumor tissue differentiation based on MRSI data [153]. Our hNMF method was introduced in Chapter 4 for MP-MRI based tumor characterization [225]. In the current chapter, an adapted version of the hNMF algorithm is considered, making it computationally more efficient while maintaining similar performance. The hNMF method consists of 2 levels of NMF, using aHALS NMF at both levels. First, rank-2 NMF is applied to the input matrix X, resulting in 2 sources ($W_{:,1}$ and $W_{:,2}$) and corresponding abundance maps ($H_{1,:}$ and $H_{2,:}$). It is assumed that each tissue class is mainly represented by one source, such that we can assign each tissue type to either $W_{:,1}$ or $W_{:,2}$. The voxels are then divided over the 2 sources based on their abundance values $H_{1,:}$ and $H_{2,:}$. Whereas hNMF originally used iterative thresholding for dividing the voxels, the algorithm has been adapted such that k-means clustering is applied to the $H_{1,:}$ and $H_{2,:}$ values for efficient voxel assignment. The clusters are initialized at [1 0] and [0 1], i.e. we expect to find one cluster (X_1) with high $H_{1,:}$ and low $H_{2,:}$ values, and another cluster (X_2) with low $H_{1,:}$ and high $H_{2,:}$ values. By doing so, we were able to

make the hNMF computation approximately 10 times faster while maintaining segmentation performance. A second level of NMF is then applied to X_1 and X_2 separately, with the rank set to the number of tissue types represented by each cluster, k_1 and k_2. k_1 and k_2 are determined based on visual inspection of the abundance maps $H_{1,:}$ and $H_{2,:}$ in combination with T1c: each tissue class is assigned to the source for which it has the highest abundance values. The sources found at this second level are the actual tissue signatures. Finally, we recombine all the voxels and use non-negative least squares (NNLS) fitting with the tissue signatures to obtain abundance maps for each tissue type over the full ROI. A schematic overview of the adapted hNMF procedure is shown in Figure 6.1. The original hNMF algorithm is presented in Figure 4.2 on page 70.

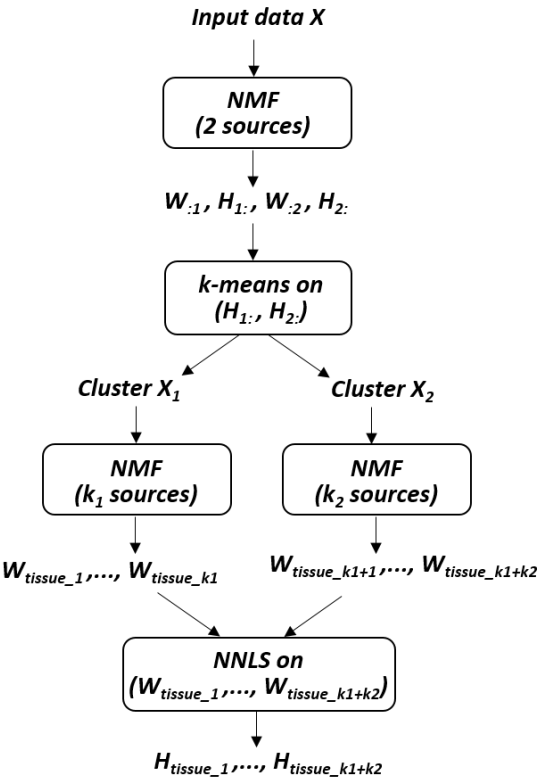

Figure 6.1: Schematic overview of the adapted hNMF algorithm.

6.3.2 Clustering

Fuzzy c-means clustering

FCM is a clustering method that allows each data point to belong to multiple clusters with varying degrees of membership. These membership grades indicate to which degree the data points belong to each cluster. FCM aims at minimizing the following objective function:

$$\underset{m,C}{\operatorname{argmin}} \sum_{i=1}^{n} \sum_{j=1}^{k} m_{i,j}^{p} \parallel x_i - c_j \parallel^2 \tag{6.2}$$

where p is the fuzziness exponent which determines the level of fuzziness, $m_{i,j}$ is the degree of membership of data point x_i to cluster C_j. Here, the data points are the columns of the input matrix X, representing the voxels' MP-MRI feature vectors. The cluster centroids c_j and the degrees of membership $m_{i,j}$ are calculated by:

$$c_j = \frac{\sum_{i=1}^{n} m_{i,j}^{p} x_i}{\sum_{i=1}^{n} m_{i,j}^{p}} \tag{6.3}$$

$$m_{i,j} = \frac{1}{\sum_{l=1}^{k} \left(\frac{\parallel x_i - c_j \parallel}{\parallel x_i - c_l \parallel} \right)^{\frac{2}{p-1}}} \tag{6.4}$$

FCM is carried out through an iterative optimization of the objective function in Equation 6.2, with the update of cluster centroids c_j and cluster memberships $m_{i,j}$ through equations 6.3 and 6.4, respectively. In the current study we have set the fuzziness exponent p to 2. Hard segmentation is obtained from FCM by assigning each data point to the cluster for which it has the highest membership value.

Gaussian mixture modeling

GMM aims at finding the maximum likelihood parameters of a mixture of Gaussians fitting the input data. This comes down to maximizing the posterior probability of a parameter set Θ of k Gaussian components, given the input data X:

$$p(\Theta|X) = \sum_{i=1}^{k} \varphi_i \mathcal{N}(X|\mu_i, \sigma_i) \tag{6.5}$$

with φ_i the weight of the ith Gaussian component and $\mathcal{N}(x|\mu_i, \sigma_i)$ the probability of a data point x (i.e. columns of the data matrix X) belonging to a normal distribution with mean μ_i and standard deviation σ_i. The Expectation Maximization (EM) algorithm [60] is the most popular technique to solve this optimization problem. EM alternates between an expectation step (E-step) and a maximization step (M-step) until convergence. In the E-step an estimation of the posterior probability $p(\Theta|X)$ is computed given the current estimation of the model parameters. In the M-step a maximum likelihood update of the model parameters is performed based on the posterior probability computed in the E-step. GMM is a soft clustering technique, i.e. it will provide probabilities of the voxels belonging to the different tissue classes. Hard segmentation is obtained by assigning each voxel to the tissue class with the highest probability. In order to recover from singular covariance matrices, a regularization constant of $1e^{-3}$ was added to the diagonal elements of the covariance matrices [65].

Spectral clustering

Spectral clustering (SC) has become a popular tool for image segmentation. Several studies have applied SC for automated brain tumor and lesion detection [163, 281]. SC refers to a family of graph partitioning methods which make use of the eigenvalue decomposition of a graph Laplacian matrix for dimensionality reduction prior to the actual clustering step [262]. Its success is mainly based on the fact that SC does not make any strong assumptions about the cluster shape nor about the density distribution of the data points. The aim of SC is to find a partitioning of the similarity graph such that the edge weights between the data clusters are very low. The graph Laplacian matrix L is derived from the adjacency matrix A, which defines the edge weights between the graph nodes. We used a k-nearest neighbor graph to obtain a sparse adjacency matrix, for efficiently performing the eigenvalue decomposition. The Gaussian similarity function $exp(-\|x_i - x_j\|^2/(2\sigma^2))$, where x_i and x_j stand for columns of the data matrix X, was used to calculate the weights in A [262]. Careful selection of the connectivity parameters is crucial in constructing the adjacency matrix, such that the connected components are well separable. We set the number of k nearest neighbors to 50, being in the order of the logarithm of the number of data points for both the UZ Leuven and UZ Gent datasets [26]. For the Gaussian similarity function, σ was set to 0.5, being in the order of the mean distance of the data points to their k^{th}-nearest neighbor [262]. Several graph

Laplacian matrices have been proposed in literature. We used the normalized Laplacian proposed by Shi and Malik [232]:

$$L = I - D^{-1}A \tag{6.6}$$

where I is the identity matrix and D is a diagonal matrix with its diagonal elements equal to the sums of the rows of A. Eigenvalue decomposition is then applied to L and a matrix U is created with the first k eigenvectors as its columns. Finally, k-means clustering is applied to the rows of U to obtain the k clusters.

6.3.3 Initialization

Except for SC, all of the proposed unsupervised classification methods are non-convex, i.e. the obtained solution will be a local rather than a global optimum, depending on the initialization conditions. A wide range of initialization methods have been developed for NMF and clustering. As NMF and clustering assume different data models, their most common initialization strategies (besides random initialization) also differ. To avoid favouring either methods, we have performed segmentation using 2 initialization strategies: the successive projection algorithm (SPA), which has been suggested for NMF source detection (see Chapter 5), and kmeans++, a popular initialization method for clustering.

SPA is a forward selection method which minimizes collinearity of the selected variables in vector space [2]. It has been commonly used as an endmember extraction tool for hyperspectral unmixing. SPA aims at finding the vertices of a convex hull spanning the dataset, under the assumptions of near-separable NMF [80]. SPA iteratively adds data points to the set of initial sources, at each step selecting the data point with highest l_2-norm in the orthogonal subspace of the already selected sources. See Section 5.3.3 for a detailed description of the SPA algorithm.

Kmeans++ is a recently proposed method for initializing the cluster centroids [4]. The intuition behind this approach is that it is favourable to spread out the initial centroids. The first cluster center is chosen randomly from the data points, after which each subsequent centroid is chosen from the remaining data points with probability proportional to its squared distance from the point's closest existing cluster center. As kmeans++ is not a deterministic algorithm, we successively ran it 20 times and withheld the result with the best (= highest or lowest) objective function value for the considered classification method. SC does not require any initialization, but we applied both initialization methods to the k-means clustering step in the final part of the algorithm.

6.3.4 Validation

The unsupervised classification methods were applied to each patient's MP-MRI dataset individually. Segmentation results obtained from the NMF and clustering methods were validated based on manual expert labelling. Manual delineations of the pathological tissue regions were obtained at the 2 hospitals from an experienced radiologist, each of them having more than 5 years of experience in neuro-radiology. The number of tissue classes present within each patient's ROI was determined based on visual inspection by the radiologists. The Dice score (see Section 2.2.1) was used to quantify the spatial alignment between each method's automated segmentation and the manual segmentation. Dice scores were calculated for the main pathological tissue types: actively proliferating tumor (active tumor), necrosis and edema. Furthermore, to be consistent with the results reported for the BRATS challenge [168], Dice scores were also reported for the tumor core (active tumor + necrosis) and the whole tumor (tumor core + edema). Average computation times were calculated for the different methods to compare their computational cost. We also wanted to assess how well the built-in modeling assumptions of the different methods align with the actual data in feature space, e.g. in terms of spatial distribution, the shape of the tissue clusters and the overlap between them. Principal component analysis was applied to the pathological data of each patient and the data points were projected onto the 2 main principal components to allow for visual inspection of the data distribution.

6.4 Results

6.4.1 UZ Gent

A comparison of the segmentation results of the unsupervised algorithms on the UZ Gent dataset is given in Table 6.2 for kmeans++ initialization and in Table 6.3 for SPA initialization. The highest mean Dice score is marked in bold for each tissue class. With both initialization methods, hNMF has the highest mean Dice score for all tissue classes except for necrosis, for which SC performs best. We also report for each method the number of times it did not detect a pathological tissue class which was annotated by the radiologist. The number of cases with undetected tissue classes ranged from 0 to 1 out of 21 patients for active tumor, from 0 to 3 out of 12 patients for necrosis and from 2 to 7 out of 15 patients for edema. The high number of cases of undetected edema is also reflected in its low Dice scores. This is mainly attributed to the relatively high number of grade III patients in the UZ Gent dataset, many of them exhibiting

non-enhancing tumor that is hard to differentiate from the peri-tumoral edema. Neither initialization method gave clearly better performance for any of the segmentation algorithms. For edema, we see that kmeans++ achieves Dice scores which are at least as high as with SPA and the number of undetected edema cases are also generally lower.

		NMF			Clustering		
		aHALS	**Convex**	**hNMF**	**FCM**	**GMM**	**SC**
Dice [%]	**Tumor**	66 ± 13	61 ± 22	$\mathbf{70 \pm 14}$	61 ± 21	66 ± 19	69 ± 14
	Necrosis	56 ± 28	54 ± 31	62 ± 25	41 ± 33	58 ± 32	$\mathbf{66 \pm 33}$
	Edema	33 ± 24	37 ± 27	$\mathbf{46 \pm 22}$	38 ± 23	36 ± 26	43 ± 24
	Core	76 ± 11	73 ± 15	$\mathbf{79 \pm 12}$	71 ± 19	77 ± 18	76 ± 16
	Whole	78 ± 13	82 ± 10	$\mathbf{86 \pm 8}$	78 ± 17	85 ± 10	85 ± 10
Undetected [#]	**Tumor**	0/21	1/21	0/21	0/21	0/21	0/21
	Necrosis	1/12	2/12	0/12	3/12	2/12	2/12
	Edema	4/15	4/15	2/15	3/15	3/15	3/15

Table 6.2: Segmentation results for the UZ Gent dataset when using kmeans++ initialization. Mean Dice score ± standard deviation is reported for active tumor, necrosis, edema, the tumor core and the whole tumor. The number of undetected cases is reported for active tumor, necrosis and edema.

		NMF			Clustering		
		aHALS	**Convex**	**hNMF**	**FCM**	**GMM**	**SC**
Dice [%]	**Tumor**	65 ± 13	64 ± 18	$\mathbf{71 \pm 15}$	60 ± 21	66 ± 21	68 ± 14
	Necrosis	55 ± 27	60 ± 28	60 ± 28	41 ± 34	48 ± 33	$\mathbf{65 \pm 29}$
	Edema	28 ± 23	18 ± 18	$\mathbf{46 \pm 23}$	38 ± 23	27 ± 26	43 ± 24
	Core	76 ± 11	74 ± 14	$\mathbf{79 \pm 12}$	70 ± 18	77 ± 18	76 ± 15
	Whole	78 ± 12	84 ± 10	$\mathbf{86 \pm 8}$	77 ± 15	84 ± 10	85 ± 11
Undetected [#]	**Tumor**	0/21	0/21	0/21	0/21	0/21	0/21
	Necrosis	1/12	1/12	1/12	3/12	3/12	1/12
	Edema	5/15	7/15	2/15	3/15	6/15	3/15

Table 6.3: Segmentation results for the UZ Gent dataset when using SPA initialization. Mean Dice scores ± standard deviation are reported for active tumor, necrosis, edema, the tumor core and the whole tumor. The number of undetected cases is reported for active tumor, necrosis and edema.

Figure 6.2 gives an example of some of the MP-MRI input maps and the segmentation results for a grade III oligo-astrocytoma patient from UZ Gent. No necrotic region was present but the main challenge was the differentiation of non-enhancing tumor from edema. Comparing with the manual segmentation, all classification methods overestimate the active tumor region. hNMF gives the best segmentation for active tumor, slightly better than aHALS NMF and Convex NMF. All methods obtain a good estimation of the whole tumor region.

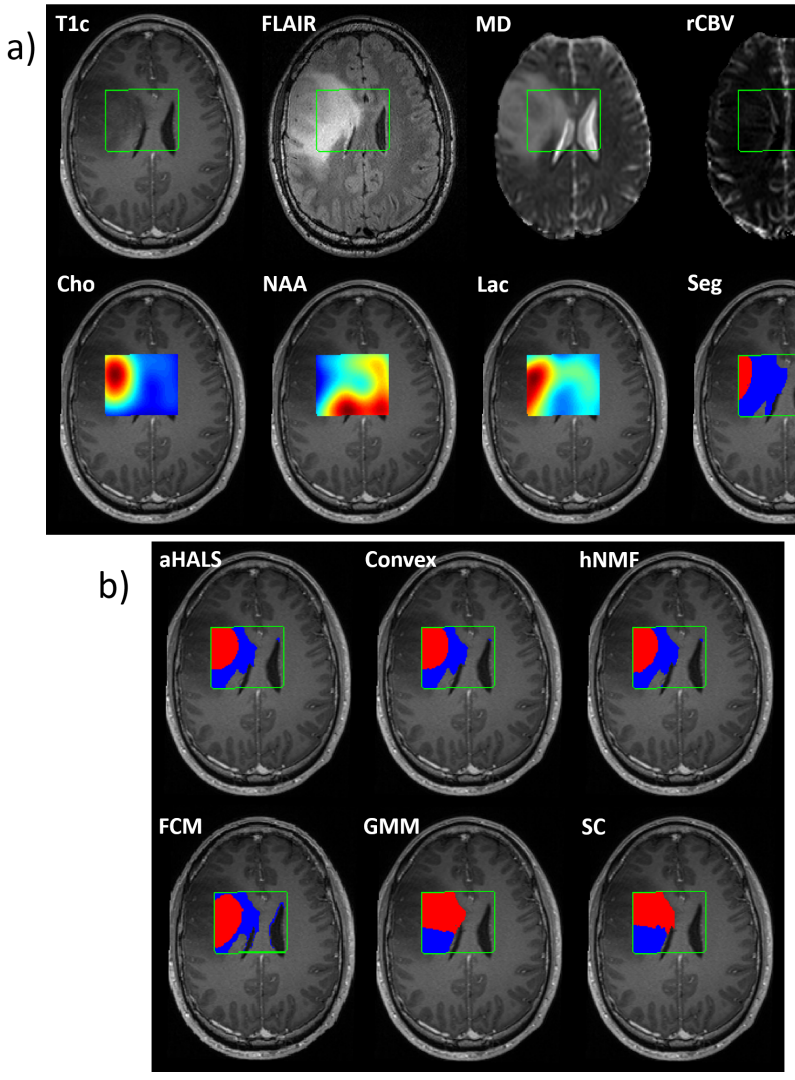

Figure 6.2: a) Some of the MP-MRI images of a grade III oligo-astrocytoma patient from the UZ Gent dataset. 'Seg' indicates manual segmentation by a radiologist (active tumor=red, edema=blue). The ROI is delineated in green. b) Segmentation results (using kmeans++ initialization) of the NMF methods (top row) and the clustering methods (bottom row).

6.4.2 UZ Leuven

The segmentation results for the UZ Leuven dataset are shown in Table 6.4 for kmeans++ initialization and in Table 6.5 for SPA initialization. With both initialization methods, hNMF has the highest mean Dice score for all tissue classes except for the tumor core, for which Convex NMF has a slightly higher mean Dice using kmeans++ initialization. aHALS NMF matches hNMF with equal mean Dice score for active tumor using kmeans++, aHALS NMF and SC match hNMF performance for the tumor core using SPA and SC matches hNMF for necrosis with both initialization methods. Globally, GMM has the worst performance on this dataset, with the lowest mean Dice in 6 out of 10 classes from both tables. Neither kmeans++ nor SPA clearly showed better performance for any of the unsupervised methods or for any of the tissue classes. The number of undetected cases ranged from 0 to 2 out of 14 patients for active tumor, from 0 to 3 out of 9 patients for necrosis and from 0 to 2 out of 9 patients for edema. hNMF matches the lowest number of undetected cases for all tissue classes, except for necrosis with SPA initialization, where Convex NMF has no undetected cases and hNMF has one.

		NMF			**Clustering**		
		aHALS	**Convex**	**hNMF**	**FCM**	**GMM**	**SC**
Dice [%]	**Tumor**	$\mathbf{73 \pm 16}$	67 ± 26	$\mathbf{73 \pm 15}$	66 ± 26	57 ± 26	72 ± 17
	Necrosis	56 ± 28	49 ± 29	$\mathbf{57 \pm 29}$	45 ± 31	35 ± 28	$\mathbf{57 \pm 35}$
	Edema	51 ± 23	60 ± 16	$\mathbf{62 \pm 21}$	60 ± 15	50 ± 29	56 ± 20
	Core	84 ± 9	$\mathbf{85 \pm 9}$	84 ± 9	83 ± 12	82 ± 10	84 ± 9
	Whole	80 ± 12	83 ± 11	$\mathbf{85 \pm 11}$	82 ± 10	83 ± 10	83 ± 11
Undetected [#]	**Tumor**	0/14	1/14	0/14	1/14	2/14	0/14
	Necrosis	1/9	2/9	1/9	2/9	3/9	2/9
	Edema	1/9	0/9	0/9	0/9	2/9	0/9

Table 6.4: Segmentation results for the UZ Leuven dataset when using kmeans++ initialization. Mean Dice scores ± standard deviation are reported for active tumor, necrosis, edema, the tumor core and the whole tumor. The number of undetected cases is reported for active tumor, necrosis and edema.

		NMF			Clustering		
		aHALS	**Convex**	**hNMF**	**FCM**	**GMM**	**SC**
Dice [%]	**Tumor**	68 ± 24	72 ± 15	$\mathbf{73 \pm 15}$	66 ± 26	63 ± 16	72 ± 17
	Necrosis	55 ± 28	55 ± 17	$\mathbf{57 \pm 29}$	45 ± 31	33 ± 27	$\mathbf{57 \pm 35}$
	Edema	50 ± 23	40 ± 28	$\mathbf{62 \pm 21}$	60 ± 15	52 ± 22	56 ± 20
	Core	$\mathbf{84 \pm 9}$	80 ± 11	$\mathbf{84 \pm 9}$	82 ± 12	81 ± 9	$\mathbf{84 \pm 10}$
	Whole	79 ± 13	82 ± 11	$\mathbf{85 \pm 11}$	82 ± 10	82 ± 9	84 ± 11
Undetected [#]	**Tumor**	1/14	0/14	0/14	1/14	0/14	0/14
	Necrosis	1/9	0/9	1/9	2/9	3/9	2/9
	Edema	1/9	2/9	0/9	0/9	1/9	0/9

Table 6.5: Segmentation results for the UZ Leuven dataset when using SPA initialization. Mean Dice scores ± standard deviation are reported for active tumor, necrosis, edema, the tumor core and the whole tumor. The number of undetected cases is reported for active tumor, necrosis and edema.

6.4.3 Computational cost

Table 6.6 reports the average computation times per patient for the UZ Gent dataset using kmeans++ initialization. These were the computationally most expensive analyses, as the UZ Gent data were 3D (UZ Leuven dataset has 2D ROI) and the analyses are repeated 20 times with kmeans++ to come to the final result. Segmentations were computed in Matlab R2015a (The MathWorks, Natick, MA, USA) with an Intel Core i7-3720QM processor and 8GB of RAM. The input data matrix typically consisted of approximately 30,000 voxels by 24 MRI features. aHALS NMF and in particular FCM clustering, both known for their computational efficiency, have relatively low computational costs. SC has a computational cost similar to aHALS NMF as we have defined a sparsely connected k-nearest neighbors graph, allowing efficient calculations. Computation times for hNMF and GMM are 2 to 4 times higher compared to aHALS NMF and SC. For hNMF, this is partly explained by its hierarchical structure, requiring 3 aHALS NMF analyses at each run, but the main expense comes from the NNLS recombining step to calculate the abundance maps at the final stage of the algorithm. Convex NMF has the highest computation time, approximately twice as high as for hNMF, due to its low convergence rate.

	Average computation time [sec]
aHALS NMF	80
Convex NMF	429
hNMF	208
FCM	31
GMM	266
SC	74

Table 6.6: Average computation time (in seconds) per patient on the UZ Gent dataset using kmeans++ initialization.

6.4.4 Data distribution

To gain some insight into the spatial distribution of the data points in feature space, the shape of the tissue clusters and their overlap, we performed a principal component analysis on the active tumor and necrosis data points of a single GBM patient from each hospital. The projections of the data points onto the first and second principal component are plotted against each other in Figure 6.3, providing insight into the data spread along the 2 main directions of data variance. We can see that the first principal component separates tumor and necrosis, but there is a distinct difference in the data spread of both patients. For the UZ Leuven patient, there is a structured pattern in the data and distinct curves of data points can be observed. Such patterns are not seen in the data of the UZ Gent patient, for which data points appear more randomly distributed. The structure in the UZ Leuven data can be explained by the limited number of voxels within the 2D ROI in combination with the large MRSI voxel size. Often only a few MRSI voxels fully cover the tumor and the necrotic tissue class. The MRSI data have been upsampled to the spatial resolution of $1 \times 1 \times 3\text{mm}^3$, generating a high number of interpolated and noiseless data points. The structure that can be seen in the data are interpolation patterns within the highly upsampled MRSI data. This explains why SC performs very well and almost as good as hNMF on the UZ Leuven data: it can easily model this kind of structured data as the SC similarity matrix is based on a distance metric among the nearest neighbors of each data point. It also explains why GMM underperforms on the UZ Leuven dataset, as its assumptions of a Gaussian within-class data distribution and ellipsoidal cluster shape does not hold. For the UZ Gent data, the structured upsampling effect is not relevant as we are analyzing a sufficiently large (3D) ROI, with about 10 times more data points compared to the UZ Leuven patients.

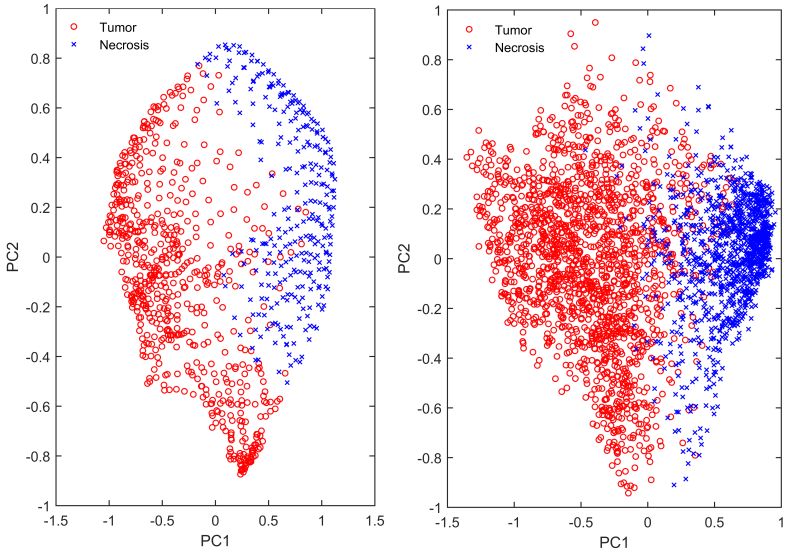

Figure 6.3: Active tumor and necrosis datapoints projected onto the plane formed by the first and second principal component for an UZ Leuven GBM patient (left) and for an UZ Gent GBM patient (right).

6.5 Discussion

6.5.1 NMF vs clustering

So far only few studies have considered MP-MRI for brain tumor segmentation, although it has been commonly suggested in literature [14, 63, 115]. In Chapter 4, it was shown that significantly higher segmentation accuracy could be achieved by applying hNMF to MP-MRI data compared to using only cMRI data [225]. For the whole tumor region, segmentation algorithms already achieve Dice scores in the range of inter-rater variability, which has been reported to be between 80 and 85% [168]. Such Dice scores were also achieved by most algorithms in the current study. Lower Dice scores are commonly reported for the individual pathological tissue types. There are several reasons why it is particularly difficult to differentiate the tumor subregions. First of all, HGGs show a high degree of within-tissue heterogeneity. Clustering methods explain the data variability in this way, by incorporating within-class variability into their model. On the other hand, the pathological region contains a continuum of tissue mixtures

rather than absolute tissue boundaries: e.g. active tumor in HGGs contains foci of microscopic necrosis [16] and edema will often be invaded by active tumor cells [203]. These tissue mixtures are further enhanced at the voxel level due to the partial volume effect. Clinical MRI is often acquired in a highly anisotropic way, with relatively low inter-slice resolution. Furthermore, advanced MRI modalities usually have lower spatial resolution than cMRI, causing even more partial volume effects. NMF has this concept of class mixtures built into its model. As such, clustering methods and NMF methods each model only one aspect of the data variability: clustering does not incorporate tissue mixtures and NMF does not consider within-tissue variability. In the current study we did not see either type of methods outperforming the other, suggesting that both aspects of data variability play a considerable role in glioma segmentation.

6.5.2 hNMF performance

Segmentation performance was compared among the different methods by reporting mean Dice scores for 5 tissue classes, on 2 independent MP-MRI datasets, using 2 different initialization methods. The hNMF method globally performed best, with the highest mean Dice score in 17 out of the 20 comparisons. hNMF also provided robust segmentation, with relatively low standard deviations and the lowest number of undetected cases for nearly all tissue classes. The hierarchical structure of hNMF gives it some advantages over the other methods. At the first (rank-2) NMF level, the tissue classes are divided over 2 sources, with tissues that are most similar in terms of MP-MRI feature vectors being assigned to the same source. At the second level, NMF can focus on the more subtle differences between the similar tissue classes. This property has allowed hNMF to better separate edema from non-enhancing tumor in the UZ Gent dataset, resulting in the highest mean Dice scores and the lowest number of undetected sources for edema in Tables 6.2 and 6.3. Another challenging problem for most unsupervised methods is how to cope with data imbalance. Unsupervised methods have to model a large amount of data with a limited number of degrees of freedom, based on a global optimization criterion. This does not favour the detection of small lesions. In the current study this was mainly an issue in detecting small necrotic regions. As the input data are divided into 2 groups after the first NMF level, hNMF might be able to detect small tissue regions more easily at the second NMF level. This has allowed hNMF to detect some small necrotic areas that other methods were not able to differentiate. An important drawback of the original hNMF algorithm as presented in Chapter 4 was its computational expense. This was mainly attributed to the iterative mask selection procedure after the first NMF level, using hard thresholding to subdivide the voxels over both sources. We have

adapted the hNMF algorithm, using k-means clustering after the first NMF level for the voxel assignment. This has brought the computation time of hNMF within the range of the other methods, as shown in Table 6.6.

6.5.3 Spatial regularization

Many classification based segmentation algorithms incorporate some kind of spatial regularization to obtain spatially consistent results. It is to be expected that most neighboring voxels will belong to the same tissue class, except at the tissue boundaries. Common techniques to encourage spatially consistent segmentation include Markov random fields [169] and spatial smoothness regularization on the NMF abundance matrix H [280]. We did not apply spatial regularization to any of the unsupervised methods, as it would bias the comparison to the other methods. In general, we did obtain spatially consistent segmentations, as is illustrated in Figure 6.2. This is partly explained by the fact that we combined many MP-MRI features, thereby increasing the specificity of the feature set. Secondly, the inclusion of locally averaged MRI features further improved robustness of the segmentation in terms of spatial coherence. Zikic *et al.* also reported naturally smooth segmentation results by combining cMRI with DWI and including spatial and context-aware features [292]. Nevertheless, unsupervised segmentation methods might benefit from adding spatial regularization, especially when reducing the number of MRI features or extending the region of interest to full brain slices.

6.5.4 UZ Gent and UZ Leuven datasets

For the UZ Gent dataset edema is clearly showing the lowest Dice scores. 8 out of 21 patients in the UZ Gent dataset (6 grade III and 2 GBM) exhibited non-enhancing tumor in combination with vasogenic edema in the ROI. Distinction of both tissue types is often difficult on conventional MRI due to the diffuse infiltrative nature of HGGs, lacking a well-defined margin. The peri-lesional environment often contains not only T2/FLAIR hyperintense vasogenic edema but also tumor components which appear less hyperintense on T2/FLAIR weighted images due to hypercellularity. The differences in signal intensity on T2/FLAIR can be subtle making the segmentation in these cases susceptible to subjective interpretation. Even when using MP-MRI, the tissue signatures of both tissue types are similar and proper differentiation remains challenging (see Figure 6.2). The UZ Leuven dataset only contained 3 (out of 14) grade III patients, and none of the patients showed non-enhancing tumor with edema in the ROI. Due to this clear bias across the patient cohorts, higher Dice scores

were found for edema in the UZ Leuven dataset. Looking at the distribution of the data points of the pathological tissue classes as shown in Figure 6.3, we did find a marked difference between both datasets. A structured pattern is seen in the UZ Leuven data, which is explained by the limited extent of the 2D ROI in combination with the low spatial resolution of MRSI compared to the other MRI modalities. SC, which is known to cope well with complex cluster shapes, performs well on the UZ Leuven data, almost as good as hNMF. GMM, due to its strong assumption of Gaussian within-class data distribution, is not able to model the structured data well and has the lowest Dice scores. However, for the 3D ROI of the UZ Gent dataset, these structured patterns are no longer appearing. Therefore, SC and GMM perform comparably to most of the other methods on the UZ Gent dataset.

6.5.5 Initialization methods

We have performed tissue segmentation using 2 initialization methods: kmeans++ which is commonly used for clustering and SPA which is commonly used for NMF. No bias was seen in the results with both types of initialization: kmeans++ did not improve segmentation results for the clustering methods nor did SPA show better results for the NMF methods. In general, similar performance was found using both types of initialization. The main difference was seen in the segmentation of edema in the UZ Gent dataset: mean Dice scores for edema were higher with kmeans++ than with SPA for 3 out of 6 classification methods, and were equal for the other 3 methods. The number of undetected cases for edema was also at least as high with SPA as with kmeans++. SPA minimizes collinearity between the initialized sources by successively selecting data points with maximal l_2-norm in the orthogonal subspace of the previously selected sources. As such, SPA might have difficulty in correctly initializing tissue classes with similar sources, such as non-enhancing tumor and edema. Indeed, we noticed that SPA often seeded only one source for the combined region of non-enhancing tumor and edema, whereas kmeans++ often obtained at least 2 sources. Kmeans++ is more likely to select data points which are further apart in Euclidian space, but it does not explicitly look for the extremal points in a dataset. As we run kmeans++ a sufficient number of times, it is likely that both non-enhancing tumor and edema get a seeding point in at least one of those runs. This does mean that we have to run kmeans++ a considerable number of times (20 times in the current study), whereas SPA guarantees reproducible results with only one run. At the individual patient level, SC, aHALS NMF and hNMF gave very similar results with both initialization methods, which is reflected in the similar mean Dice scores for most tissue classes. aHALS NMF has previously been shown to be insensitive to the initialization strategy,

and the same goes for hNMF as it consists of 2 levels of aHALS NMF [221]. SC does not require an initialization, but it applies k-means clustering to the matrix of eigenvectors in the final stage of the algorithm. This k-means step however does not seem sensitive to the initialization method, probably because the complexity of the classification problem is already reduced by withholding only the k main eigenvectors to find k tissue classes.

6.5.6 Supervised vs unsupervised classification

Recently, supervised classification methods have been receiving most attention for automated brain tumor segmentation [168]. However, unsupervised methods remain competitive, e.g. Juan-Albarracin *et al.* have recently combined some dedicated cMRI pre-processing steps with several unsupervised methods, including FCM, GMM and hidden Markov random fields. Their results ranked among the best supervised algorithms in the BRATS challenge [115]. A direct comparison between our results and those from the BRATS challenge is not in place, as we are combining cMRI data with PWI, DWI and MRSI in a multi-parametric approach, whereas BRATS only considers cMRI data. Furthermore, we restricted our study to the limited region of interest of the MRSI data, whereas full 3D imaging data of the brain are considered in the BRATS challenge. Nevertheless, several of the unsupervised algorithms reported in this study, and in particular hNMF achieved mean Dice scores which are higher than the best performing algorithms participating in the BRATS 2012-2013 challenge [168], showing the potential of these methods. Unsupervised methods are more flexible as they don't require an extensive set of training data, which is only valid for segmentation under a particular data acquisition and processing protocol. This limitation might be more restrictive when considering MP-MRI data, as there is a wide variability and fast evolution in acquisition protocols and post-processing methods for PWI, MRSI and DWI [283]. Furthermore, unsupervised methods are not susceptible to overfitting, and they don't require intensity calibration, which can be challenging in the presence of pathological tissue [14]. On the other hand, an attractive feature of supervised methods is that they don't require a rank estimation step for determining the number of tissue types, and they automatically apply voxel labelling based on the training data. The unsupervised methods presented in this study are not fully automatic: they require the user to specify the number of tissue types to be found within the region of interest. Therefore, further research will focus on automation by incorporating prior knowledge, e.g. tissue probability maps based on an image atlas [115], feature-specific knowledge [119] or user input [168].

6.6 Conclusion

In this chapter, we explored the usage of unsupervised classification methods applied to MP-MRI data for segmenting the tumor subregions in HGGs. We have proposed a hierarchical 2-level hNMF method and compared its performance to 5 common unsupervised classification algorithms. hNMF achieved the best segmentation results in terms of mean Dice scores on 2 independent MP-MRI datasets acquired at different hospitals. This trend was confirmed using both kmeans++ and SPA initialization. Unsupervised methods can be applied to any MP-MRI dataset at the individual patient level, irrespective of the acquisition protocol or image processing methods used. Future work should focus on further automation of unsupervised segmentation methods by incorporating prior knowledge. This will be the subject of the next chapter.

Chapter 7

Semi-automated brain tumor segmentation on MP-MRI data using regularized NMF

The NMF-based segmentation algorithms that were considered in the previous chapters are not fully automated: they require the user to specify the number of tissue types present in the region of interest, and they don't provide automatic tissue labelling. Furthermore, the analyses have been restricted to a limited region of interest around the tumor, as MRSI data were not available for the whole brain. A clinically relevant segmentation tool needs to perform well on full 3D imaging datasets. To allow for further automation and to be able to obtain competitive results, unsupervised methods rely on the usage of prior knowledge. As gliomas exhibit a wide variability in terms of size, shape, location and appearance, imposing proper prior knowledge is difficult. Incorporation of prior knowledge often limits the segmentation algorithm to a specific subset of gliomas or to a specific set of MRI features. This chapter presents a semi-automated framework for brain tumor segmentation based on regularized NMF. L1-regularization is incorporated into the NMF objective function to promote spatial consistency and sparseness of the tissue abundance maps. The pathological sources are initialized through user-defined voxel selection. Knowledge about the spatial location of the selected voxels is combined with tissue adjacency constraints in a post-processing step to enhance segmentation quality. As the user-defined prior knowledge is

patient-specific, the proposed method is applicable to any type of glioma and to any MP-MRI dataset. This study has been submitted as a journal paper for peer review as [223].

7.1 Introduction

In recent years, significant advancements have been made in the field of automated brain tumor segmentation. Both semi-automated and fully automated methods have been proposed [87]. A popular approach is to combine imaging biomarkers from different MRI sequences on a voxel-wise basis, thereby increasing specificity and reducing overlap between tissue classes. Nowadays, supervised classification methods are receiving most attention [168]. These methods rely on an extensive set of training images with manually annotated ground truth to learn decision boundaries between the tissue classes in feature space. The most popular methods include random forests [13, 292], support vector machines [147, 259] and neural networks [114]. Additional constraints may be imposed to further enhance performance. For instance, one can assume that a voxel is likely to belong to the same tissue class as its neighbors, instead of classifying all voxels independently. Markov Random Fields [152] and Conditional Random Fields [131] are common approaches to encourage such spatial consistency, either as a regularization term in the optimization function [13, 147] or as a post-processing step [245].

Unsupervised classification methods are also being considered for tumor segmentation. These methods are very flexible, as they don't require an extensive training dataset with a uniform acquisition protocol. They learn classification rules directly from the imaging data at hand, based on some similarity criterion. Popular approaches include fuzzy c-means clustering (FCM) [73, 119], Gaussian mixture modeling [115, 169], hidden Markov Random Fields [175, 288] and non-negative matrix factorization (NMF) [182, 225]. Due to the absence of training data, unsupervised methods rely more strongly on the incorporation of prior knowledge to obtain competitive results. As gliomas exhibit a wide variability in terms of size, shape, location and appearance, imposing proper prior knowledge is difficult. Prastawa *et al.* used a normal brain atlas to obtain healthy tissue priors, and added a tumor prior based on the assumption of tumor enhancement on T1c [202]. These priors along with the MRI features were then fed to a Gaussian mixture model for tissue classification. The authors acknowledged that their algorithm might not adequately detect tumors which are non-enhancing or tumors with a large mass effect. Fletcher-Heath *et al.* applied knowledge-based FCM to segment non-enhancing gliomas [73]. Clusters containing tumor as well as CSF were separated based on PD-enhancement.

The tumor region was then separated from the ventricular CSF by assuming symmetry of the ventricles across the hemispheres. Finally, extraventricular CSF and other false positives are removed from the tumor mask using morphological post-processing steps, assuming only one tumor volume. Kazerooni *et al.* applied FCM combined with manually initiated region growing to segment GBMs on a MP-MRI dataset combining T2, DWI and PWI [119]. The active tumor region was found by applying FCM to the PWI data, assuming hyperperfusion throughout the region. Saha *et al.* presented a computationally efficient tumor detection algorithm, by assuming symmetry across the hemispheres in healthy brain [218]. A bounding box is obtained within one hemisphere which maximizes the histogram-based dissimilarity with the corresponding contralateral region based on the Bhattacharya coefficient. Görlitz *et al.* applied a hidden Markov Random Field model to MRSI data of brain tumors [88]. They used a graph-cut algorithm to efficiently solve the spatio-spectral model, resulting in a binary tumor mask without differentiating the tumor subcompartments. Similarly, Ortega-Martorell *et al.* applied Convex NMF to MRSI data of brain tumor-bearing mice to delineate the entire pathological brain area [182]. Evidently, many unsupervised segmentation studies incorporate prior knowledge in the form of simplifying rules, limiting their applicability to a subset of gliomas with specific characteristics: the tumor is assumed to be enhancing or non-enhancing, deformation of the surrounding healthy tissue is assumed to be small, multi-focal lesions can not be detected, the lesion does not extend to the contralateral hemisphere, etc. Some methods are limited to a particular set of MRI parameters, as they rely on one or more specific imaging biomarkers. Other methods simplify the segmentation problem by only providing an outer contour of the pathological region, without differentiating active tumor from the other tumor subcompartments.

Semi-automated segmentation algorithms use a different approach. They incorporate prior knowledge in the form of user-specific input, either as an initialization or as a post-processing step. Competitive performance has been reported for semi-automated algorithms [168]. Kwon *et al.* combined a normal brain atlas with a tumor growth model, based on user-defined tumor seeds [129]. This allowed them to generate a patient-specific atlas which also models the tumor mass effect. Through the user-defined seeding, multi-focal gliomas could be handled as well. Hamamci *et al.* initialized a cellular automata (CA) algorithm based on the maximal tumor diameter as drawn by the user on T1c images [96]. An active level-set surface is then initialized from the user-defined maximal diameter, to impose spatial smoothness on the segmented pathological region. The authors acknowledged the limited applicability of their method to enhancing tumors. Havaei *et al.* initialized a k-nearest neighbors (kNN) classifier with user-selected voxels in the tumor subcompartments as well as in the healthy brain regions [99]. In addition to the MRI imaging

features, spatial coordinates of the seeding points were also exploited in the feature set. Spatial consistency was encouraged by adding regularization in the form of Markov Random Fields and Conditional Random Fields, which further improved segmentation results. The main drawback of semi-automated methods is that they don't provide reproducible results, as the segmentation depends on subjective user input. Robustness against user input variability is therefore a key aspect of these methods. In this chapter, a semi-automated tumor segmentation method is proposed based on regularized NMF. User-defined seeding points in the pathological regions are combined with a sophisticated seeding method for the normal brain tissues to initialize the NMF algorithm. Piece-wise spatial smoothness as well as sparseness of the NMF tissue abundance maps are encouraged through L1-regularization. Morphological post-processing based on the spatial location of the user-defined seeds is exploited to further remove false positives. The proposed method is applicable to all types of gliomas with any MP-MRI dataset, as the user-defined prior knowledge is patient-specific. To verify robustness against user input variability, each patient's dataset is repeatedly analyzed with randomly selected seeding points from the pathological subcompartments.

7.2 MP-MRI dataset

The HGG patient cohort of UZ Gent was considered for this study (9 grade III gliomas, 12 GBMs). cMRI, PWI and DWI data were included in the MP-MRI feature set and were pre-processed as described in Section 3.2. MRSI data were omitted from the analyses, to extend segmentation to full image slices. Table 7.1 gives a concise overview of the considered MRI sequences and the resulting MP-MRI parameters. Six MRI features were obtained from the raw acquired data after pre-processing: T1, T1c, FLAIR, rCBV, MD and b0. All MP-MRI features were coregistered and resampled to the same spatial resolution of $1 \times 1 \times 3\text{mm}^3$. cMRI data were skull-stripped and T1c served as a reference for rigid co-registration, using the normalized mutual information criterion [161] and cubic B-spline interpolation for reslicing. Analyses were performed on 10 axial slices located around the tumor centre. Additional intensity-based features were added to the feature set to include localized spatial information. An in-plane local neighborhood of 3×3 and 5×5 voxels was used to calculate average intensity values that were assigned to the central voxel. Each feature's full range was linearly rescaled to [0-1].

	UZ Gent dataset
cMRI	**T1, T1c, FLAIR**
PWI	*Gradient echo DSC-MRI* **rCBV**
DWI	*Spin echo DWI (b_0, b_{500}, b_{1000} in resp. 1, 3, 3 directions)* **MD, b0**

Table 7.1: Schematic overview of the MP-MRI protocols and derived parameters of the UZ Gent dataset.

7.3 Methodology

7.3.1 Regularized NMF

The most commonly used objective function to measure the quality of the NMF factor approximation is the Frobenius norm, which is based on the squared Euclidean distance:

$$\min_{W,H} f(W,H) = \min_{W,H} \quad \frac{1}{2}\|X - WH\|_F^2 \quad ,\forall i,j : W_{i,j} \geq 0, \quad H_{i,j} \geq 0 \qquad (7.1)$$

Many algorithms have been developed to solve the above NP-hard NMF formulation [267]. In this chapter, structured data fusion NMF (SDF NMF) is being used as implemented in Tensorlab [260]. A transformation of variables is used to convert the constrained optimization problem in Equation 7.1 into an unconstrained problem, by squaring the entries of the factor matrices. The Gauss-Newton algorithm with dogleg trust region (GNDL) is used to solve the resulting non-linear least-squares problem [177]. At each iteration of the GNDL algorithm, a step is calculated by iteratively solving a linearized version of Equation 7.1 using conjugate gradients. The SDF framework allows straightforward incorporation of regularization terms into the objective function. An L1-regularization term was included to promote both spatial consistency as well as sparseness of the tissue abundance maps:

$$\min_{W,H} f(W,H) = \min_{W,H} \quad \frac{1}{2}(\|X - WH\|_F^2 + \lambda\|H(L+I)\|_1) \qquad (7.2)$$

The L1-regularization term in Equation 7.2 in fact has 2 components. The first component, HL, promotes piece-wise smoothness on the tissue abundance maps.

L is a sparse $n \times n$ matrix, with each row containing a vectorized Laplacian kernel. As such, each row of L applies a two-dimensional second order spatial derivative to the corresponding voxel. An in-plane neighborhood of 4 voxels is considered for the Laplacian kernel. Due to the relatively low out-of-plane resolution of the MRI data, neighboring voxels in adjacent slices are not considered. The second component applies L1-regularization to H directly, imposing sparseness to the abundance maps. A maximum number of 500 iterations and a convergence tolerance of 10^{-6} for the relative difference of two subsequent values of the objective function were used as stopping criteria for the NMF analyses. The regularization coefficient λ was empirically set to 0.1.

7.3.2 NMF initialization

As NMF poses a non-convex optimization problem, initialization of the factor matrices (W_0 and H_0) is required. Careful initialization is important, as it might speed up the convergence process and influence the final segmentation result. The pathological tissue sources are initialized based on voxel selection. The user is required to select one or more voxels in each pathological region, i.e. active tumor and, if present, necrosis and edema. For each selected voxel, a candidate source vector is added to W_0 as the averaged feature vector of the selected voxel and its 4 in-plane neighboring voxels. For each pathological tissue class, we then calculate the correlation coefficient between the candidate source vectors as their inner vector product after normalization. The initial pathological sources are obtained after merging candidate source vectors with a correlation coefficient higher than 0.95 in W_0 and replacing them by their average, thereby reducing complexity of the NMF model.

As the normal brain tissue types (i.e. white matter, gray matter, CSF and blood vessels) are still highly abundant in the affected brain, an automated seeding procedure is used to obtain their initial source vectors. To cope with the variance within the tissue classes, 2 sources are assigned to each normal tissue type, leading to a total of 8 normal tissue sources. We use the Successive projection algorithm (SPA) [2] to obtain an initial estimate of the normal sources. SPA returns a subset of voxels from the input matrix X with minimal collinearity in feature space (see Section 5.3.3 for a detailed description of the SPA algorithm). Similarly to the pathological sources, the normal sources are obtained as the averaged feature vector of each selected voxel and its 4 in-plane neighboring voxels, assuming these to belong to the same tissue type as the selected voxel. As SPA is known to be sensitive to outliers [80], an additional FCM procedure is applied, using the initial pathological sources and the normal sources obtained from SPA to initialize the cluster centroids. FCM alternatingly updates the cluster centroids and the cluster membership values (see Section

6.3.2 for a detailed description of the FCM algorithm). As we already obtained proper initialization of the pathological tissue types through user input, the pathological cluster centroids are forced to remain the same throughout the FCM updating procedure. The final FCM centroids make up the columns of W_0. The columns of H_0 are found by applying non-negative least squares fitting to the corresponding column of X and W_0. To reduce computation time, a downsampling factor of 2 was applied to the x- and y-direction of each image slice. After NMF analysis, the obtained tissue abundance maps were upsampled again through linear interpolation.

7.3.3 Morphological post-processing

After NMF analysis, each voxel is assigned to the source for which it has the highest abundance value. Voxels that were assigned to different sources belonging to the same pathological tissue class are merged to obtain a preliminary tissue segmentation mask. A morphological post-processing procedure, consisting of 2 steps is then applied to the pathological segmentation masks to further remove false positives, by exploiting the known location of the user-selected voxels. Spatial consistency of the tissue regions is assumed, therefore the segmentation masks are analyzed in terms of (3D-)connected components. For each selected pathological tissue type, only the connected component closest to each user-defined voxel is withheld in the corresponding tissue segmentation mask (Step 1). As the user-defined voxels are not necessarily selected close to the centre of the tissue region, the distance between a selected voxel and a connected component is defined as the minimal distance between the selected voxel and any voxel from the connected component. It is possible that the same connected component is found for several or all user-selected voxels of the same tissue class.

To avoid missing disjoint regions of a tumor component for which no voxels were selected by the user, an additional step is performed which assumes spatial connectivity of the various pathological components (Step 2). Necrotic areas are always adjacent to an active tumor region. Therefore, any connected component of the preliminary necrosis mask which has a common edge with any withheld active tumor component is also included. Similarly, components of the preliminary active tumor mask are also included if they have a common edge with any withheld necrotic component. The same procedure is also applied to edema, by verifying spatial adjacency to the active tumor region(s). Figure 7.1 illustrates the morphological post-processing procedure for the necrotic tissue mask. One small necrotic component is not withheld after Step 1, as the user didn't click in this region. However, it was still recovered in Step 2 by verifying spatial connectivity with the active tumor mask.

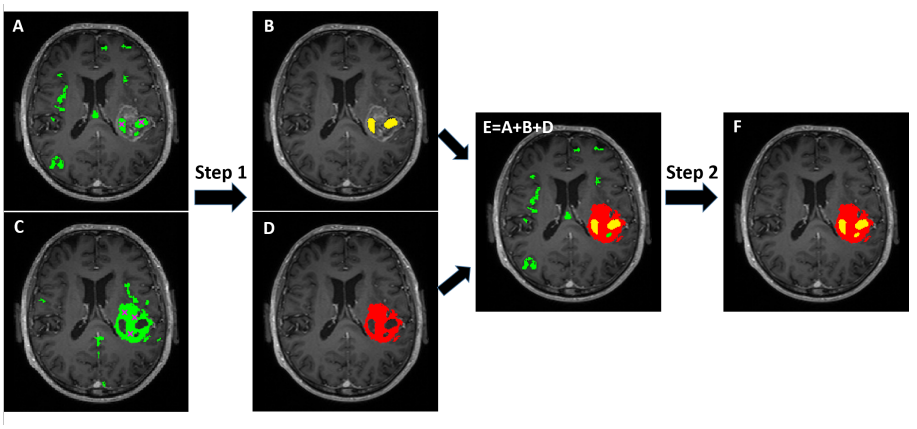

Figure 7.1: Illustration of morphological post-processing after initial semi-automated NMF based segmentation of necrosis (A) and active tumor (C). Step 1: false positives are removed by withholding only the connected components closest to the user-defined seeding points (marked by an 'x' in A and C). Step 2: spatial adjacency of the connected components in the preliminary necrosis mask (green) to the withheld active tumor mask (red) is verified in E. The final necrosis mask is shown in yellow in F.

7.3.4 Validation

In analogy to previous work [168], Dice scores are reported for active tumor, the tumor core (active tumor + necrosis) and the whole tumor (tumor core + edema). To verify robustness of the semi-automated method to user input variability, the NMF analysis was repeated 20 times per patient with different selection of the seeding points. In each run, an automatic random selection of 3 points was performed in each pathological region, based on the manual segmentation masks. Besides NMF with spatial regularization and sparseness ($\mathbf{NMF}_{\text{spatial_sparse}}$), additional NMF analyses were performed to assess the added value of the regularization term and the advanced MRI modalities: NMF without regularization ($\mathbf{NMF}_{\text{no_reg}}$), NMF with only spatial regularization ($\mathbf{NMF}_{\text{spatial}}$), and NMF with spatial regularization and sparseness when only considering cMRI data ($\mathbf{NMF}_{\text{cMRI}}$). For the $\mathbf{NMF}_{\text{cMRI}}$ analyses, rCBV and MD were omitted from the MP-MRI feature set but b0 was withheld, acting as a surrogate for T_2-weighted MRI.

7.4 Results

Figure 7.2 gives an example of the regularized $\mathbf{NMF}_{\text{spatial_sparse}}$ results and tissue segmentation for a GBM patient with a typical ring-enhancing lesion. Some of the MP-MRI input maps are shown on the first row. The NMF abundance maps as well as the final segmentation masks are shown for the pathological tissue types on the second row.

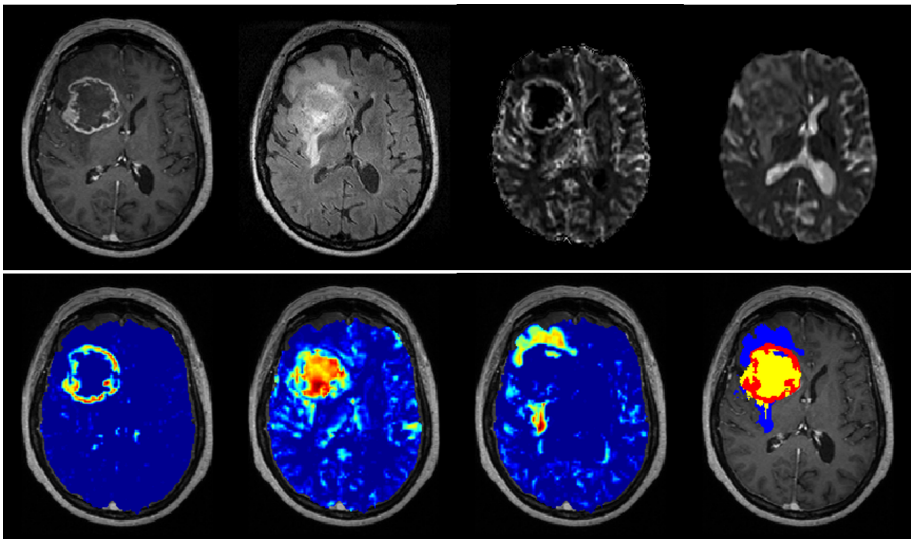

Figure 7.2: First row: coregistered MP-MRI maps of a GBM (patient 3), left to right: T1c, FLAIR, rCBV, MD. Second row, left to right: NMF abundance maps for active tumor, necrosis and edema. The final segmentation masks are shown on the right for active tumor (red), necrosis (yellow) and edema (blue).

Figure 7.3 shows the dispersion of the Dice scores for active tumor per patient over 20 runs. 16 out of 21 patients have a median Dice score of at least 60%. 15 patients have a lower quartile Dice score of at least 50%. 17 patients have an interquartile range not higher than 15%. 7 patients had at least one run with a Dice score lower than 30%. A large spread in the Dice scores is seen for patient 6, who has a non-enhancing anaplastic astrocytoma with focal areas of enhancing GBM. Variability in the results was caused by the difficult differentiation of non-enhancing tumor from edema. Patient 8 suffers from a non-enhancing bifocal GBM, with only a small enhancing area in the smaller lesion. Due to the random selection of voxels in the active tumor region, in some runs voxels were only selected from one of both lesions, resulting in a failure to detect the other lesion. A large dispersion in the Dice scores is also found for patient 18,

exhibiting a heterogeneous and irregularly shaped GBM with varying degrees of enhancement in the active tumor region. Active tumor and necrosis were not well differentiated in some runs, mainly due to ambiguous voxel selection near the pathological tissue boundaries. The low outlier scores for patient 10 are due to the difficult differentiation of non-enhancing tumor from edema. The low outlier score for patient 13 is explained by a suboptimal and unrepresentative random voxel selection in the active tumor region.

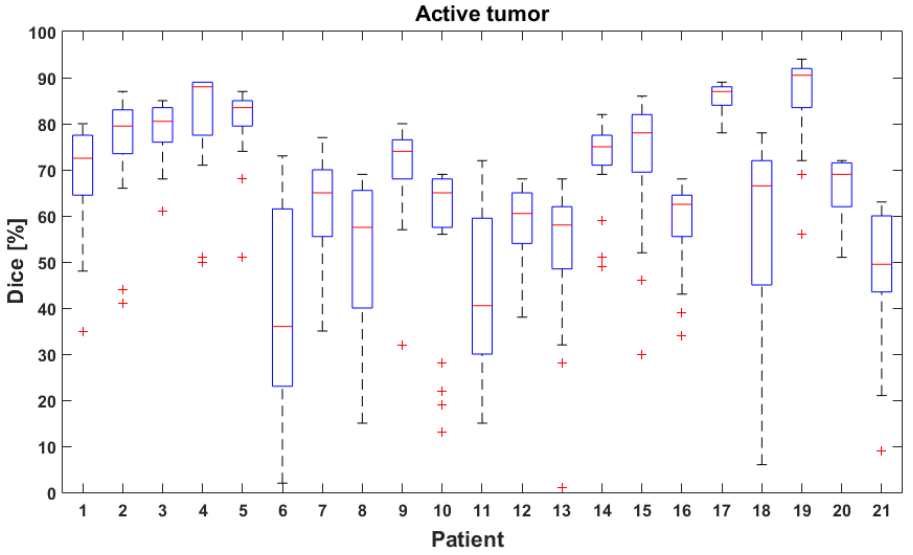

Figure 7.3: Boxplots showing the dispersion of the Dice scores for active tumor. Boxplots show quartile ranges of the Dice scores, '+' indicates outliers.

Boxplots for the tumor core are shown in Figure 7.4. 18 out of 21 patients have a median Dice score higher than 65%. For 17 patients, the lower quartile Dice score was at least 60%. The interquartile range was not higher than 15% for 17 patients. 7 patients have at least one run with a Dice score lower than 40%. For 5 out of the 7 patients, these low Dice scores were found to be outliers (i.e. at a distance of more than 1.5 times the interquartile range from the lower quartile value).

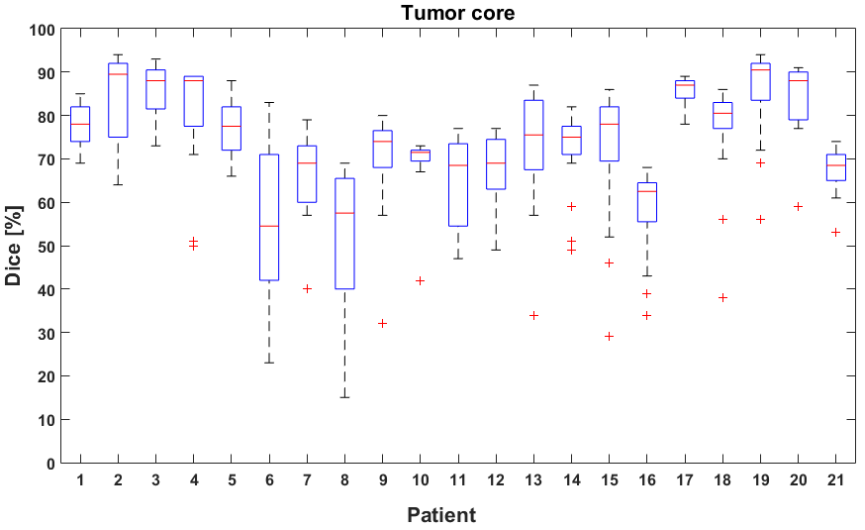

Figure 7.4: Boxplots showing the dispersion of the Dice scores for the tumor core. Boxplots show quartile ranges of the Dice scores, '+' indicates outliers.

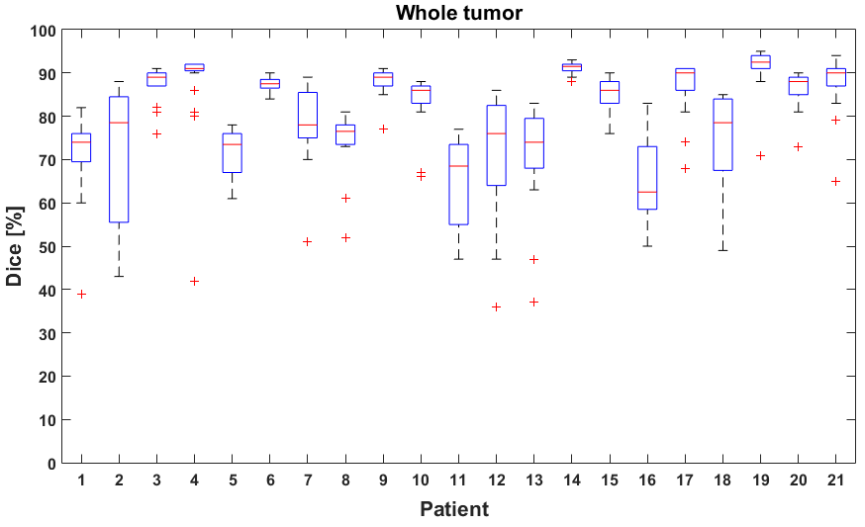

Figure 7.5: Boxplots showing the dispersion of the Dice scores for the whole tumor region. Boxplots show quartile ranges of the Dice scores, '+' indicates outliers.

Figure 7.5 shows the boxplots for the whole tumor region. A median Dice score higher than 70% was found for 19 out of 21 patients. 18 patients have lower quartile value higher than 60%. 5 patients have at least one run with a Dice score lower than 45%. For 4 out of these 5 patients, such low values were found to be outliers.

Table 7.2 gives a comparison of the mean Dice scores over all patients for the different NMF analyses ($\mathbf{NMF}_{\text{no_reg}}$, $\mathbf{NMF}_{\text{spatial}}$, $\mathbf{NMF}_{\text{spatial_sparse}}$ and $\mathbf{NMF}_{\text{cMRI}}$). Overall, the best performance is obtained with $\mathbf{NMF}_{\text{spatial_sparse}}$. When considering only spatial regularization ($\mathbf{NMF}_{\text{spatial}}$), segmentation results are mainly worse for the active tumor region, where the Dice scores are even lower than for $\mathbf{NMF}_{\text{no_reg}}$. The spatial smoothing was found to be too severe for several GBMs with a narrow ring-enhancing active tumor compartment. Compared to $\mathbf{NMF}_{\text{no_reg}}$, Dice scores are higher with $\mathbf{NMF}_{\text{spatial_sparse}}$ for active tumor and the tumor core, with an increase of 2% in most cases. The Dice score for the whole tumor region is 0 to 1% higher for $\mathbf{NMF}_{\text{spatial_sparse}}$ compared to $\mathbf{NMF}_{\text{no_reg}}$. $\mathbf{NMF}_{\text{spatial_sparse}}$ does not show any improvement compared to $\mathbf{NMF}_{\text{cMRI}}$ for the active tumor region. An increase in Dice score of 1 to 2% is found for the tumor core and an increase of 3 to 4% for the whole tumor region.

	Active tumor			**Tumor core**			**Whole tumor**		
	25_{th} prentile	**Mean**	**75_{th} prentile**	**25_{th} prentile**	**Mean**	**75_{th} prentile**	**25_{th} prentile**	**Mean**	**75_{th} prentile**
$\mathbf{NMF}_{\text{no_reg}}$	58	63	72	66	72	79	76	80	84
$\mathbf{NMF}_{\text{spatial}}$	55	60	70	65	71	79	76	80	85
$\mathbf{NMF}_{\text{spatial_sparse}}$	60	65	74	68	74	80	77	80	85
$\mathbf{NMF}_{\text{cMRI}}$	60	65	74	67	72	79	73	77	82

Table 7.2: Mean Dice scores [%] for NMF without regularization, with spatial regularization (λ=0.1), with spatial regularization and sparseness (λ=0.1), and using only cMRI data with spatial regularization and sparseness. For each tissue class, the first column reports the mean of the 25th percentile Dice score across all patients, the second column the mean of the mean Dice score across all patients and the third column the mean of the 75th percentile Dice score.

7.5 Discussion

Most of the unsupervised brain tumor segmentation algorithms in literature incorporate prior knowledge in the form of some basic assumptions or heuristics, which often limits their general applicability to a subset of gliomas or to a particular set of MP-MRI data. Prior knowledge in the form of user input is flexible, allowing to incorporate patient-specific information regarding appearance, location and/or shape of the tumor. Using our semi-automated framework, we were able to tackle the main limitations of many unsupervised classification methods like NMF. Most unsupervised algorithms require proper initialization to come to a valid locally optimal solution. Several studies cope with this by running the segmentation algorithm numerous times with a randomized initialization, then selecting the final solution based on a predefined objective function [115, 225]. We are combining user-defined voxel selection for initializing the pathological sources with a sophisticated initialization of the normal brain tissue sources based on SPA and FCM. As SPA is sensitive to outliers, its output is fed into the FCM algorithm, thereby providing a deterministic initialization for FCM.

Due to the lack of labelled training data, automatic assignment of a tissue label to each segmented region is non-trivial for unsupervised methods. Several studies do not explicitly propose an automated labelling strategy [169, 175], while others assume specific characteristics of the tumor compartments in specific MRI images (such as e.g. contrast enhancement of active tumor) [73, 119]. In our proposed methodology, tissue labelling automatically results from the voxel selection in the pathological regions.

Unsupervised methods model the input with a limited number of degrees of freedom, based on a global optimization criterion. As such, they tend to obtain data clusters of similar size and they do not promote the detection of small pathological tissue regions [127]. Some studies cope with this by using an over-determined model, i.e. by setting the number of clusters/sources higher than the actual number of tissue classes. This approach also allows the detection of regions with different characteristics belonging to the same tissue class (e.g. enhancing and non-enhancing tumor). However, it requires an intelligent post-processing for fusing the clusters from the same tissue class. Juan-Albarracin *et al.* applied hierarchical agglomerative clustering to merge similar classes using the Jensen-Shannon divergence criterion to measure distances between the probability density functions of different clusters [115]. Hsieh *et al.* obtained 32 segmented regions after FCM combined with region growing, then eliminated non-pathological areas and merged pathological areas based on assumptions regarding tumor size, symmetry across the hemispheres and spatial adjacency [104]. We used on average 12 to 13 sources to model 6 to 7 tissue classes, with

the number of sources depending on the number of initial pathological sources that were merged. Fusing of the segmented regions was straightforward, as pathological regions had already been labelled through voxel selection.

False positive regions were excluded from the pathological tissue masks by exploiting the spatial location of the selected voxels in a morphological post-processing procedure. As the pathological tissue classes are assumed to form spatially consistent regions, only the connected component closest to each selected voxel is withheld in the segmentation masks. Loss of true positive tumor components not selected by the user is avoided by assuming spatial adjacency of the pathological regions. Similar types of post-processing have been proposed to enhance segmentation results. Cordier *et al.* only withheld 1 or 2 connected components for each tissue region based on their size, assuming that the largest component(s) correspond to the true tissue region [56]. Menze *et al.* exploit knowledge about shape and location of the tumor based on an atlas tissue prior to remove false positives [169]. Bauer *et al.* apply spatial regularization in the form of a conditional random field, penalizing neighboring voxels with different labels depending on the likelihood of tissue adjacencies [13]. Havaei *et al.* initialize a k-nearest neighbors (kNN) classifier based on voxel selection by the user [99]. They do not exploit the spatial location of the selected voxels in a post-processing step, but instead they add voxel coordinates to the feature set of the kNN classifier.

Comparison to other segmentation studies is sometimes hampered by the fact that different segmentation metrics are being used, or because of a different definition of pathological subregions. We have quantified our results for active tumor, the tumor core and the whole tumor region using the Dice score, in accordance with the Multimodal Brain Tumor Segmentation (BRATS) challenge held at the Medical Image Computing and Computer Assisted Intervention (MICCAI) conference. Looking at the results reported for HGGs on the BRATS 2012 and 2013 datasets [168], it can be seen that our results are competitive, even when looking at the mean of 25th percentile values (see Table 7.2). For the whole tumor region, Dice scores are close to or even higher than 80%, which is in the range of inter-observer variability [168]. Overfitting to the presented dataset seems unlikely as we are using an unsupervised method with no tuning parameters, except the weight of the regularization term. Excellent results compared to literature have been reported for other semi-automated segmentation methods as well, either based on a supervised [99] or an unsupervised classifier [119, 129]. It has to be noted that the current BRATS dataset only contains cMRI data. Comparison might therefore be more appropriate to the $\mathbf{NMF}_{\text{cMRI}}$ results, for which only cMRI data were considered. The inclusion of additional MRI modalities for brain tumor segmentation has been commonly suggested [14, 39] and initial studies have found improved

segmentation results with extended MP-MRI datasets [64, 225]. When including MD and rCBV into the MRI feature set, we found no advantage for the active tumor region, but Dice scores increased by 1 to 2% for the tumor core and by 3 to 4% for the whole tumor region. These findings are in accordance with [225], where inclusion of PWI and DWI were found to be mainly advantageous for the tumor core and the whole tumor region.

Nowadays, a common approach to further improve segmentation results is to model the spatial dependency between the tissue labels of adjacent voxels, typically using a Markov Random Field (MRF) [99, 169] or a Conditional Random Field (CRF) [13, 147]. MRF and CRF are usually combined with a probabilistic classifier which segments all voxels independently. With MRF, a fixed penalty term is added to the cost function when adjacent voxels have different labels. This penalty term is also depending on the (dis)similarity of the feature vectors when using CRF. To the author's knowledge, this is the first study to consider spatially regularized NMF for tumor segmentation. We have added a spatial regularization term with sparseness to the NMF objective function. L1-regularization was used to promote piece-wise smoothness of the abundance maps, allowing for discontinuities at the tissue boundaries. Whereas MRF and CRF are directly applied to the voxel labels, we applied L1-regularization to the tissue abundance maps which represent continuous variables. In some patients, suboptimal segmentation results were found when using spatial regularization only. This is reflected in the lower Dice scores for $\mathbf{NMF}_{\text{spatial}}$ in the active tumor region and in the tumor core (Table 7.2). The original abundance maps have a relatively low degree of sparseness, such that piece-wise spatial smoothness was not always restricted to the true tissue boundaries. Results were improved by combining spatial regularization with sparseness. This kind of regularization resembles MRF more closely, since MRF implies absolute sparseness of the tissue labels. Over all patients, we found an increase in mean Dice score by 2% for active tumor and for the tumor core when using spatial regularization with sparseness (Table 7.2). For the whole tumor region, an increase of 1% was found for the 25th and 75th percentile Dice scores and no increase was found for the mean value. Similar improvements in segmentation results have been reported in the literature [99, 147].

As imaging characteristics vary considerably across glioma subtypes and grades, supervised classification methods require extensive training datasets with a uniform acquisition protocol. Separate classifiers are usually built for LGGs and HGGs to increase specificity [168]. Using training data from different patients to segment a new test case requires careful calibration of image intensities across different patients, which is non-trivial in the presence of pathology. Our method uses labelled data only from the patient under study. Feature vectors averaged over the local neighborhood of the selected voxels are assumed to be

approximative prototypes of the pathological tissue regions to be segmented.

So far, most semi-automated segmentation studies did not quantify the effect of user input variability on their results. We have assessed user input variability by repeatedly selecting random voxels in each pathological region. Fairly robust segmentation results were found for most patients (see Figures 7.3, 7.4 and 7.5), but a large spread in the Dice scores and low outlier values were found in several patients as well. In some cases suboptimal results could be explained by ambiguous voxel selection near the tissue boundaries. Figure 7.6 illustrates a case of suboptimal voxel selection for patient 1. One of the necrotic voxels was selected at the edge of the manually segmented necrotic region (red voxel in B, other necrotic voxels were selected on other slices). This voxel is located very near the active tumor region and does not show the hypo-intense T1c signal which is characteristic for necrosis. The resulting NMF segmentation is shown in C for active tumor and in D for necrosis. The necrotic voxel selected at the boundary results in an overestimation of the necrotic region and an underestimation of the tumor region. The low outlier Dice score for patient 1 for active tumor in Figure 7.3 corresponds to this voxel selection. Another patient had a bi-focal tumor, and in some runs only one of both lesions was detected as no voxels were selected in the other lesion. To obtain more consistent segmentation results, we require the user to select voxels in a smart way. Spatially distributing the selected voxels is one way of covering intra-tissue heterogeneity. As our method exploits spatial information, at least one voxel should be selected in each lesion in case of multi-focal lesions. A clinician is more likely to select voxels which are clearly located in one tissue region than voxels near diffuse tissue boundaries. Careful voxel selection is expected to further improve segmentation results and reduce variability.

7.6 Conclusion

We have presented a semi-automated brain tumor segmentation method, based on NMF with L1-regularization to promote spatial consistency and sparseness of the tissue abundance maps. The semi-automated framework was applied to an MP-MRI dataset consisting of cMRI, PWI and DWI. User-defined voxel selection is applied to initialize pathological sources of the NMF analysis, and to exploit knowledge about the spatial location of the tumor to remove false positives in a post-processing step. In this way, we aimed at incorporating prior knowledge while maintaining general applicability of our method to any type of glioma and to any MP-MRI dataset. Sensitivity to user input variability was explored through repeated analyses with different voxel selection. Robust results were found for most patients, although careful voxel selection is mandatory

to avoid sub-optimal segmentation. The reported segmentation scores are competitive with state-of-the-art in literature, even when looking at the lower quartile values across patients.

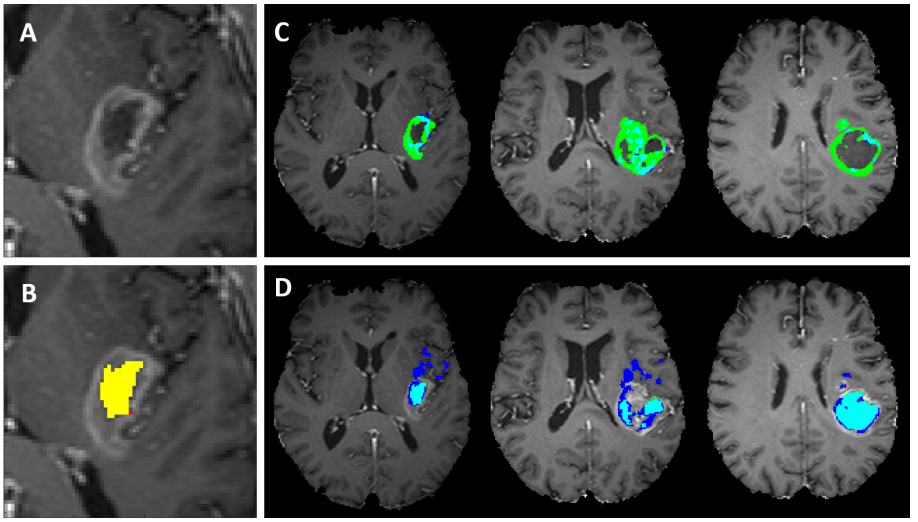

Figure 7.6: A close-up of a GBM lesion on an axial T1c slice (A). The manually segmented necrotic region is shown in yellow (B), the selected necrotic voxel is marked in red. Segmentation of the active tumor region (C) and necrotic region (D) on several slices, blue indicates NMF segmentation, green indicates manual segmentation and cyan indicates overlap.

Chapter 8

Application of semi-automated regularized NMF to the BRATS 2013 Leaderboard dataset

In the previous chapter, a semi-automated NMF-based method for brain tumor segmentation has been introduced. Competitive segmentation results were reported on the UZ Gent dataset. However, only qualitative comparison to other segmentation methods from literature could be made, as these methods have been applied to other brain tumor datasets. The Multimodal Brain Tumor Segmentation (BRATS) challenge is the first initiative to provide a publicly available MRI dataset of brain tumor patients with ground-truth segmentation, allowing for an objective comparison between segmentation algorithms. In this chapter, the NMF-based semi-automated segmentation framework is applied to the BRATS 2013 Leaderboard dataset. The dataset consists of multi-contrast cMRI scans of 21 HGG and 4 LGG patients. Bias field correction is applied to the cMRI data as a pre-processing step, to compensate for magnetic field inhomogeneities. Initialization of the pathological sources is based on voxel selection by an experienced radiologist. This voxel selection is repeated several times, to be able to quantify intra-user variability of the segmentation results. The study has been accepted for publication in the proceedings of the 2016 International Conference on Signal Image Technology & Internet-based Systems [222].

8.1 Introduction

Due to the lack of a publicly available brain tumor MRI database with ground-truth segmentations, so far most segmentation studies have evaluated their algorithms on relatively small private datasets. This has complicated comparison of the performance of different methods against each other in an unbiased way. Different combinations of MRI parameters are often used, and there is no consistency in the tumor subcompartments that are considered. Furthermore, varying metrics are being used for quantifying segmentation performance and differences in manual segmentation by human expert raters might further bias comparison. As a consequence, it has so far been difficult to assess which segmentation strategies may be worthwhile to pursue in clinical practice and research.

For the purpose of general comparison, Prastawa *et al.* provided simulated multi-sequence MRI data of brain tumors including an objective ground truth [201]. The simulated tumor images were obtained by combining realistic digital BrainWeb phantoms [55] with a well-defined and accepted tumor-growth model [54]. It models infiltrating edema adjacent to tumors, local distortion of healthy tissue and central contrast enhancement, combined with a routine for synthesizing texture similar to that of real MR images.

More recently, the Multimodal Brain Tumor Segmentation (BRATS) challenge has made available an open database of brain tumor images, serving as a benchmark for objectively comparing segmentation algorithms [168]. The BRATS challenge was first organized in 2012 in conjunction with the international conference on Medical Image Computing and Computer Assisted Interventions (MICCAI), and has been organized annually since then. A unique dataset of MR scans of both low- and high-grade glioma patients is provided. Ground-truth labelling is based on the consensus among several human expert raters. A training dataset with known ground-truth is made available, which can be used by different groups to tune their algorithms. Evaluation of the segmentation methods is then based on a test dataset, for which the ground-truth is not made publicly available. Participants can upload their segmentations through an online tool for evaluation. In this way, results are not influenced by overtraining of the method being tested, and they are truly representative of the method's segmentation performance in practice.

The aim of this chapter is to evaluate performance of the semi-automated L1-regularized NMF segmentation method (introduced in Chapter 7) on the BRATS benchmark. The method will be applied to the BRATS 2013 Leaderboard dataset, which is currently one of the most commonly considered benchmark datasets for brain tumor segmentation. L1-regularized NMF is initialized

through voxel selection in the tumor subcompartments by an experienced radiologist. This voxel selection is repeated several times, to be able to quantify intra-user variability. Segmentation results are compared against the algorithms submitted to the BRATS 2013 challenge as reported in [168].

8.2 BRATS 2013 Leaderboard dataset

The BRATS 2013 Leaderboard dataset consists of multi-contrast cMRI scans of 21 HGG and 4 LGG patients. Out of the 25 cases, 11 HGG and 4 LGG patients made up the test dataset of the 2012 BRATS challenge. 10 new HGG cases were added to the dataset for the 2013 challenge. The images were acquired at different hospitals over the course of several years, using MR scanners from different vendors and with different field strengths (1.5T and 3T). For each patient, T_1-weighted, T_2-weighted, FLAIR and post-contrast T_1-weighted (T1c) images were provided. All images were skull-stripped, co-registered to the T1c sequence and resampled to $1mm^3$ isotropic resolution. A rigid registration model was used with the mutual information similarity metric as implemented in ITK [106]. The following tissue classes were differentiated by expert annotation: necrosis, edema, non-enhancing tumor and enhancing tumor. Ground-truth labels were not available and segmentation results had to be submitted online for evaluation.

8.3 Methodology

8.3.1 Bias field correction

Intensity inhomogeneity is a common artefact in MRI acquisitions. These magnetic field inhomogeneities, also known as bias fields, are unavoidable effects, consisting of low-frequency intensity variations that corrupt the images. As most automated segmentation algorithms assume the same intensity level throughout the same tissue class, bias fields must be corrected for to ensure proper segmentation. Various bias field correction algorithms have been proposed [263], but the non-parametric non-uniform normalization (N3) approach formulated in [237] has established itself as a reference method. The method has the advantage of not requiring any tissue model for applying the correction. Recently, Tustison *et al.* proposed a new implementation of N3 called N4, which improves the N3 algorithm with a better B-spline fitting function and a hierarchical optimization scheme for the bias field correction [251]. Some severe bias field artefacts were apparent in the cMRI Leaderboard dataset. We used the N4 method as

implemented in the Slicer software [72] for bias field correction. Figure 8.1 shows the successful removal of an inhomogeneous hyperintense region posteriorly on an axial T1c slice of a GBM patient.

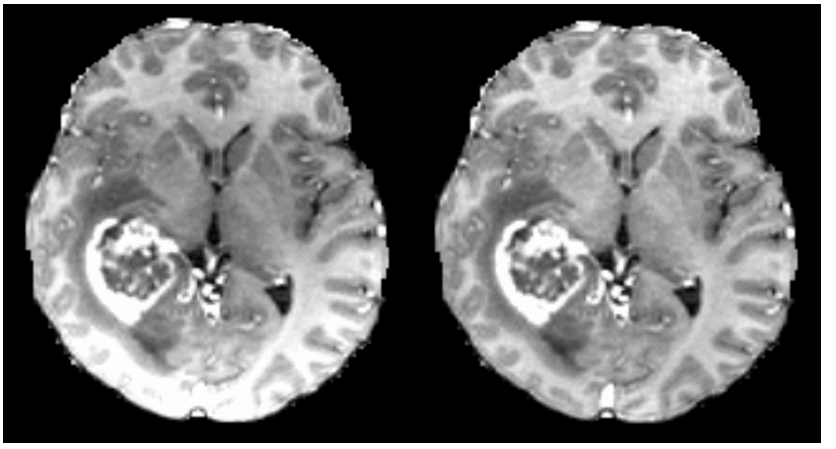

Figure 8.1: Uncorrected (left) and bias field corrected (right) axial T1c slice of HGG patient 131.

8.3.2 NMF-based semi-automated segmentation

A semi-automated tumor segmentation method based on L1-regularized NMF has been introduced in the previous chapter. The pathological NMF sources are initialized through user-defined voxel selection. The initial pathological sources are obtained after averaging candidate source vectors with a correlation coefficient higher than 0.95. The healthy tissue sources are subsequently initialized by combining SPA with FCM. To cope with intra-tissue variability, 2 sources are foreseen for each healthy tissue class. It was found in the previous chapter that the best segmentation results were obtained by including sparse and spatial regularization on the tissue abundance maps in the NMF objective function ($\mathbf{NMF}_{\text{spatial_sparse}}$):

$$\min_{W,H} f(W,H) = \min_{W,H} \quad \frac{1}{2}(\|X - WH\|_F^2 + \lambda\|H(L+I)\|_1) \tag{8.1}$$

The L1-regularization term in Equation 8.1 in fact consists of 2 components: the first component promotes piece-wise spatial smoothness, whereas the second component promotes sparseness of the tissue abundance maps. The constrained

optimization problem is solved using structured data fusion (SDF) NMF as implemented in Tensorlab [260] (see Section 5.3). After NMF analysis, each voxel is assigned to the tissue class for which it has the highest abundance value. A morphological post-processing procedure consisting of 2 steps is then applied to further enhance segmentation results. In the first step, only the connected components of the pathological segmentation masks are withheld in which at least one voxel was selected by the user. To recover true positive regions in which no voxels were selected, known spatial adjacency constraints of the pathological subcompartments are exploited in a second step. For a more detailed discussion of the semi-automated segmentation framework, the reader is referred to Section 7.3.

As in Chapter 7, the regularization coefficient λ was set to 0.1. A maximum number of 500 iterations and a relative convergence tolerance of 10^{-6} were used as stopping criteria for the NMF analyses. To initialize the pathological sources, an experienced radiologist (MA, UZ Gent) selected voxels in the tumor subcompartments of each patient. In accordance with the BRATS definition of the pathological tissue classes, the active tumor class was subdivided into a non-enhancing and an enhancing tumor class, and voxels were separately selected in both regions. The number of voxels selected by the radiologist in the tumor subcompartments (i.e. enhancing tumor, non-enhancing tumor, necrosis and edema) varied between 3 and 5. To be able to quantify intra-user variability of the segmentation results, the radiologist was asked to repeat voxel selection on the same patients 2 more times, and $\mathbf{NMF}_{\text{spatial_sparse}}$ was applied on each set of voxels. An interval of at least 24 hours was admitted between subsequent voxel selection of the same patient. For the initialization of the normal tissue sources, the FCM fuzziness exponent m (as defined in Equation 6.4) was set to 1.5, as a lower value assumes less fuzziness, i.e. more sparse tissue abundances (the default value for m is 2). To reduce computation time, NMF analyses were limited to a slice range comprising the pathological region. Furthermore, a downsampling factor of 2 was applied to the x- and y-direction of each image slice. After NMF analysis, the obtained tissue abundance maps were upsampled again through linear interpolation.

8.3.3 Validation

The BRATS challenge reports validation scores for the following tissue classes: enhancing tumor, the tumor core (i.e. enhancing tumor, non-enhancing tumor and necrosis) and the whole tumor (i.e. the tumor core and edema). Besides the Dice score, PPV and sensitivity values (see Section 2.2) are also reported to evaluate segmentation overlap.

8.4 Results

Table 8.1 compares the performance of our $\mathbf{NMF}_{\text{spatial_sparse}}$ method (Sauwen) to the best performing methods from the BRATS 2013 challenge, as well some other methods reported in literature. As $\mathbf{NMF}_{\text{spatial_sparse}}$ was repeated 3 times on each patient using different voxel selection, averaged results over the 3 runs are reported. The 5 best performing methods of the BRATS challenge all used fully automated methods based on supervised classification: 3 of them (Tustison, Meier and Reza) were based on random forest classifiers, 1 method (Zhao) applied a Markov random field model on a supervoxel-level, and 1 method (Cordier) considered patch-based classification based on k-nearest neighbors [168]. So far, the best results reported in literature on the Leaderboard dataset are based on the method by Kwon *et al.* [129]. They combine a healthy brain atlas with a tumor growth model, using seeding points selected by the user to initialize tumor growth. The tissue priors obtained from this model are then fed into a Gaussian mixture model (GMM) along with the MRI features to obtain the final segmentation. Competitive results were also reported by Juan-Albarracin *et al.* [115]. They also applied GMM to a set of MRI features including intensity values and texture features, then used a healthy brain atlas to distinguish the actual pathological components.

As can be seen from Table 8.1, competitive segmentation results are obtained using $\mathbf{NMF}_{\text{spatial_sparse}}$. Looking at the enhancing tumor region, our method reaches a mean Dice score of 0.64, which is the highest Dice score among all considered methods. For the tumor core, we obtain a Dice score of 0.64, which is also among the best performing methods for this tissue class, although Kwon *et al.* clearly outperform the competition with a Dice score of 0.79. For the whole tumor region, we obtain a Dice score of 0.72, which is rather low as all other methods except for Meier obtain higher Dice scores. Also when looking at the PPV and sensitivity values, it can be seen that $\mathbf{NMF}_{\text{spatial_sparse}}$ is one of the best methods for the enhancing tumor region, scores quite well for the tumor core but relatively low for the whole tumor region. Unlike most other methods, $\mathbf{NMF}_{\text{spatial_sparse}}$ obtains PPV values that are higher than the corresponding sensitivity values. This suggests that $\mathbf{NMF}_{\text{spatial_sparse}}$ performance is relatively more impacted by false negatives than by false positives, compared to the other methods.

	Dice			PPV			Sensitivity		
	Enhancing	**Core**	**Whole**	**Enhancing**	**Core**	**Whole**	**Enhancing**	**Core**	**Whole**
Tustison	0.53	0.65	0.79	0.51	0.70	0.83	0.66	0.73	0.81
Zhao	0.47	0.59	0.79	0.50	0.55	0.77	0.53	0.77	0.85
Meier	0.53	0.60	0.72	0.48	0.62	0.65	0.64	0.69	0.88
Reza	0.51	0.56	0.73	0.48	0.64	0.68	0.63	0.57	0.79
Cordier	0.46	0.61	0.75	0.43	0.61	0.79	0.52	0.72	0.78
Sauwen	**0.64**	**0.64**	**0.72**	**0.71**	**0.71**	**0.78**	**0.65**	**0.69**	**0.72**
Kwon [129]	0.59	0.79	0.86	0.60	0.84	0.88	0.63	0.81	0.86
Albarracin [115]	0.60	0.59	0.74	0.60	0.55	0.71	0.66	0.71	0.81

Table 8.1: Mean Dice score, PPV and sensitivity values of the enhancing tumor, tumor core and whole tumor region on the BRATS 2013 Leaderboard dataset. The first 5 rows rank the best performing groups that participated in the BRATS 2013 challenge. Our results are reported on row 6 (Sauwen) and rows 7 and 8 show other results from literature.

To assess robustness of the segmentation results against user input variability, Table 8.2 shows the Dice score ranges per patient over the 3 repeated runs with different voxel selection. In about half of the cases (36 out of 71 cells, marked in green), the Dice score did not vary by more than 5% over the different runs. 75% of the Dice score ranges (53 out of 71 cells, green and yellow) were below 10%, and 86% of the Dice scores (61 out of 71 cells, green, yellow and orange) varied by no more than 15%. Higher Dice score variations were mainly attributable to 3 patients (HGG 166, 131 and 137), exhibiting Dice score ranges larger than 15% for each tissue class (marked in red). When considering mean Dice score variations at the group level, differences were not higher than 3% for any tissue class.

	Dice score range [min-max]		
Patient ID	Enhancing tumor	Tumor core	Whole tumor
HGG 116	0.44-0.72	0.52-0.71	0.62-0.80
HGG 117	0.01-0.02	0.21-0.23	0.79
HGG 119	0.84-0.87	0.70-0.80	0.73-0.83
HGG 120	0.81-0.83	0.83-0.86	0.85-0.86
HGG 130	0.70-0.72	0.72-0.78	0.59-0.68
HGG 131	0.43-0.79	0.31-0.53	0.50-0.69
HGG 133	0.68-0.70	0.56-0.58	0.67-0.69
HGG 134	0.57-0.65	0.49-0.56	0.64-0.69
HGG 135	0.66-0.73	0.63-0.74	0.79-0.83
HGG 136	0.01-0.14	0.39-0.56	0.61-0.74
HGG 137	0.17-0.42	0.25-0.48	0.35-0.60
HGG 201	0.66-0.72	0.59-0.66	0.55-0.61
HGG 202	0.51-0.61	0.67-0.74	0.79-0.82
HGG 203	0.78-0.90	0.77-0.88	0.79-0.87
HGG 204	0.69-0.76	0.78-0.89	0.82-0.85
HGG 205	0.78-0.82	0.86-0.89	0.67-0.77
HGG 206	0.79-0.80	0.88-0.89	0.90-0.91
HGG 207	0.74-0.78	0.87-0.88	0.86
HGG 208	0.83-0.85	0.82-0.83	0.88-0.89
HGG 209	0.71-0.77	0.72-0.84	0.77-0.86
HGG 210	0.64-0.67	0.85-0.86	0.87-0.90
LGG 103	-	0.07-0.10	0.15-0.18
LGG 105	-	0.07-0.10	0.33-0.44
LGG 109	-	0.74-0.79	0.77-0.82
LGG 116	-	0.74-0.76	0.77-0.79

Table 8.2: Dice score ranges (min-max) for the individual patients over 3 repeated runs using different voxel selection. Color coding: green = range<5%, yellow = range<10%, orange = range<15% and red = range>15%.

Figure 8.2 gives an illustration of the initial segmentation of all the tissue sources after L1-regularized NMF (middle), and the final segmentation of the tumor subcompartments after morphological post-processing (right). As can be seen, some false positive regions of edema were segmented in gray matter areas of the frontal cortex.

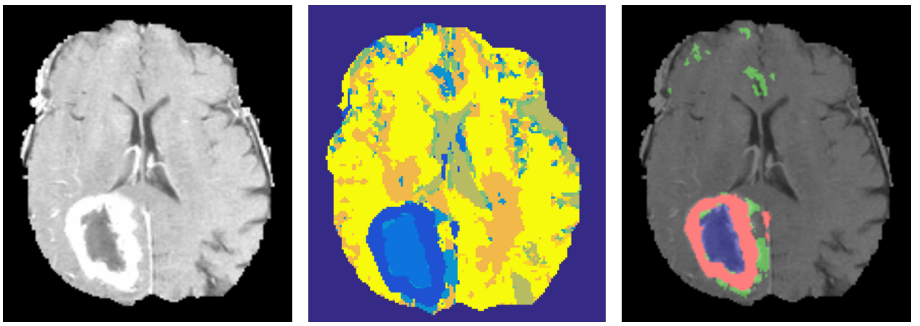

Figure 8.2: Left to right: Axial T1c slice of HGG patient 205, initial segmentation of all tissue sources after semi-automated NMF, and the final segmentation of the tumor subcompartments after morphological post-processing (red=enhancing tumor, blue=necrosis and green=edema).

8.5 Discussion

Although the vast majority of segmentation algorithms participating in the BRATS 2012 and 2013 challenges were based on supervised classifiers, it has been shown in this chapter that unsupervised classifiers incorporating adequate prior knowledge can be highly competitive. In fact, the best segmentation results that have so far been reported on the BRATS 2013 Leaderboard dataset were obtained by Kwon *et al.*, combining tissue prior probabilities obtained from a tumor growth model with GMM [129]. Besides a healthy brain atlas on which the tumor growth model is based, additional prior knowledge in the form of user input was included: user-defined seeding points were used along with an approximate radius estimation of the lesion(s) to initialize tumor class parameters. In recent years, the inclusion of atlas-based tissue priors into a discriminative segmentation method (often GMM) has become one of the most popular and successful approaches of unsupervised segmentation of brain tumors [67, 94, 169]. Accuracy of tissue prior estimates might be reduced by severe deformations of healthy brain structures due to the tumor mass effect. This might be partially coped with by combining tissue priors with imaging based features into a subsequent classifier. Using a tumor growth model, this mass effect can actually be incorporated into the estimation of the tissue priors. Although previously, affine co-registration was used to match a tumor patient's imaging dataset with the atlas [202], non-linear registration is nowadays more commonly used to obtain more accurate voxel-wise tissue priors. An important drawback of this approach is the long computation time of non-linear registration, which can take more than one hour per patient [115].

Computation times will obviously be even longer when using a tumor growth model [129].

Compared to the best performing segmentation methods from the BRATS 2013 challenge, as well as some relevant methods reported in literature, competitive segmentation results were obtained using $\mathbf{NMF}_{\text{spatial_sparse}}$. A mean Dice score of 0.64 was found for the enhancing tumor region. To our knowledge, no method reported in literature has so far reached such a high mean Dice score for enhancing tumor on the Leaderboard dataset. With a mean Dice score of 0.64, $\mathbf{NMF}_{\text{spatial_sparse}}$ also ranks among the better methods for the tumor core. For the whole tumor region, a mean Dice score of 0.72 was found, which is relatively low compared to the other algorithms: several groups reached Dice scores close to 0.80 or even higher. One of the main reasons for this fallback is the relatively high amount of false positives for edema (see e.g. Figure 8.2). Mainly, gray matter voxels exhibited feature vectors similar to edema. Gray matter structures were sometimes wrongfully annotated as edema and could not properly be removed using morphological post-processing when connected with the true edema region. Additional incorporation of prior knowledge will probably allow further improvement of the segmentation results. The usage of an atlas, for instance, could probably prevent a considerable amount of these false positives, especially at some distance from the tumor where healthy tissue deformation is negligible. As mentioned before, this would however considerably increase computation time. One could also exploit the fact that the tumor and its subcomponents usually form compact, highly connected regions. As such, a criterion might be formulated for omitting weakly connected structures from the pathological segmentation masks in the morphological post-processing step.

Looking at the results in Table 8.1, it was found that $\mathbf{NMF}_{\text{spatial_sparse}}$ has higher PPV values than sensitivity values for the same tissue class. For most other methods, this is the other way around. $\mathbf{NMF}_{\text{spatial_sparse}}$ benefits from knowledge of the spatial location of user-selected voxels, which allows removal of false positive regions that can not be applied by other methods. It seems that, compared to other methods, $\mathbf{NMF}_{\text{spatial_sparse}}$ has a relatively high share of false negatives that limit its performance. To cope with intra-tissue variability, the user is allowed to select several voxels for each pathological tissue class. Different sources are withheld for the same tissue class when they exhibit dissimilar feature vectors. The criterion to merge sources of the same tissue class is that the correlation coefficient between the sources is higher than 0.95. However, this threshold is somewhat arbitrary, and on the Leaderboard dataset it resulted in merging of all the sources in most cases. One prototype/source per tissue class might in some cases not be sufficient to adequately model within-tissue variability, and correlation coefficients higher than 0.95 were also regularly found between sources representing different tissue types. A more advanced and

adaptive way of defining the threshold for merging the sources might therefore be more appropriate, and should also take into account correlation values with respect to the other tissue classes.

Computation time of the full semi-supervised $\mathbf{NMF}_{\text{spatial_sparse}}$ procedure per patient depended on the slice range to be analysed and the rate of convergence. On average, a computation time of about 75 minutes was found. For the BRATS 2012 and 2013 challenges, reported computation times varied considerably among participants, with some algorithms taking a few minutes per patient while others required more than one hour [168], like our method. Although the treatment of brain tumors does not require immediate feedback for clinical intervention, computation times of one hour or more are inconvenient, and might not be accepted in clinical practice. The high computational cost of an algorithm is mostly attributed to a pre- or post-processing step (e.g. non-linear co-registration to an atlas or computation of advanced imaging-based features) [168]. For the semi-automated NMF-based segmentation, however, the NMF computation itself takes up most of the time. It must be noted that, compared to other common unsupervised classifiers such as GMM, NMF requires a high number of parameters to be estimated. NMF provides the relative contribution of each tissue type in each voxel. More efficient NMF algorithms do exist and might have to be considered to bring down computation time. There were mainly 2 reasons why SDF NMF was used in the current work. First of all, SDF NMF allows the incorporation of L1-regularization into the objective function. Other NMF implementations such as HALS NMF or Convex NMF do not straightforwardly incorporate L1-regularization, as it renders the NMF objective function non-differentiable [51]. Secondly, as has been shown in Chapter 5, SDF NMF is sensitive to the initialization of the factor matrices. Whereas insensitivity to initialization might be an asset in the case of suboptimal initialization, in the current work it is known that initial estimates of the pathological sources are already quite accurate, as they are based on user input. Therefore, a method such as aHALS NMF which is insensitive to the initialization might move away from the initialization instead of locally optimizing the factor matrices. To reduce computation time, in-plane downsampling of the images was applied with a factor 2 in both slice directions. Out-of-plane downsampling is not implemented, as often MRI voxels are anisotropic, with relatively low out-of-plane resolution. In case of the BRATS data, however, interpolated voxels had an isotropic resolution of $1mm^3$. As such, computation times can be further reduced by also allowing out-of-plane downsampling.

Robustness against user input variability is an important requirement that needs to be fulfilled by semi-automated methods. Intra-user variability of semi-automated $\mathbf{NMF}_{\text{spatial_sparse}}$ was verified by repeating analyses 3 times with

different voxel selection by the same radiologist. Variation of the Dice score per patient and per tissue class was found to be within 10% in 75% of the cases and within 15% in 86% of the cases. Mainly for 3 patients, 1 out 3 runs gave Dice scores for all tissue classes that were considerably lower, decreasing by close to 20% or more, compared to the other 2 runs. Mazzara *et al.* have reported an average intra-rater variability of $20 \pm 15\%$ for the manual segmentation of brain tumor volumes [165]. Therefore, a certain degree of segmentation variability is unavoidable. Nevertheless, further development of semi-automated NMF-based segmentation should not only focus on higher performance but also on increased robustness of the results.

Several groups that obtained excellent results on the BRATS 2013 challenge, such as Tustison *et al.* and Meier *et al.*, have included texture features combined with MR intensities into their feature set [168]. It has been observed that different tumor areas exhibit different textural patterns [118]. Textures can be computed according to different strategies [250]. First order statistical texture features are commonly considered. They are based on the computation of histograms in a local 2D or 3D neighborhood centred at each voxel of the image, and include for instance the mean, variance, skewness and kurtosis of these histograms. In the current chapter as well as in the previous chapters, we have only included the mean intensity over a 3×3 and 5×5 in-plane neighborhood into the feature set. Inclusion of other types of texture features has been explored, but was not found to be successful in combination with NMF. Several texture features, such as local variance and local intensity range, exhibit extreme values near tissue boundaries. For instance, a tumor voxel near the boundary with necrosis will exhibit much higher texture variance than a tumor voxel located more centrally within the enhancing region. Based on such a feature, unsupervised methods would not tend to group these 2 voxels into the same cluster. In fact, very few unsupervised classification methods in literature apply texture features. To the author's knowledge, only Juan-Albarracin *et al.* have combine unsupervised classifiers with first-order texture features (mean, skewness and kurtosis, not variance), but they did apply principal component analysis for feature reduction and it is unclear up to what extent the texture features have contributed to segmentation performance [115]. Non-linear supervised classifiers can be easily taught that voxels with a different texture behaviour still belong to the same tissue class. This seems to be a disadvantage of unsupervised over supervised methods. However, using different source vectors/prototypes for the same tissue class, it might still be possible for unsupervised methods to fully exploit texture features as well. Again, proper prior knowledge would have to be included to automatically fuse these clusters into one tissue class, for instance by only considering MR intensity features for labelling after the actual classification. Also, it must be noted that several groups, such as ourselves and in particular Kwon *et al.* obtained very good segmentation results without inclusion of

additional texture features.

8.6 Conclusion

In this chapter, we have applied semi-automated segmentation using L1-regularized NMF to the BRATS 2013 Leaderboard dataset. This has allowed to compare our method directly to the best performing methods from the BRATS 2013 challenge, as well as to several competitive methods from literature. Our semi-automated NMF-based approach outperformed all other methods in terms of mean Dice score for the enhancing tumor region. Competitive results were also found for the tumor core, but relatively low Dice scores were found for the whole tumor region. Incorporating additional prior knowledge should allow further improving segmentation quality, in particular for the whole tumor region, as well as increasing robustness of the segmentation results.

Chapter 9

Conclusions and future perspectives

The following main research questions have been addressed throughout this thesis:

- Can we use NMF-based methods to perform brain tumor segmentation and to characterize tumor heterogeneity?
- Can we improve segmentation performance by including advanced MRI modalities into the MP-MRI feature set? What is the added value of the individual MRI modalities?
- How sensitive are the segmentation results to (h)NMF initialization and how can we optimize initialization of the factor matrices?
- How do the NMF-based methods compare to other unsupervised classification methods in terms of segmentation performance?
- How can we further automate NMF-based segmentation by incorporating prior knowledge, while at the same time maintaining general applicability of the method to any MP-MRI dataset and to any type of brain tumor?
- How robust are semi-automated NMF-based segmentation results to user input variability?
- How does semi-automated NMF-based segmentation perform in direct comparison to other state-of-the-art segmentation methods?

9.1 Conclusions

In **Chapter 4**, a hierarchical NMF method was presented to characterize brain tumor heterogeneity based on MP-MRI data. hNMF basically consists of 2 levels of NMF. An iterative voxel selection procedure was used at the second level of NMF to minimize cross-contamination between the tissue sources. hNMF provides tissue-specific signatures of the pathological tissue classes, as well as abundance maps of their spatial distribution. The abundance maps were converted into absolute tissue segmentation and good spatial alignment was found with manual segmentation by a neuro-radiologist.

The UZ Leuven dataset was used in **Chapter 4** to assess the added value of the individual MRI modalities. To this purpose, hNMF analyses were also performed on reduced MP-MRI datasets, leaving out one MRI modality at a time. For the HGG patients, Dice scores were significantly higher for most tissue classes with the full MP-MRI dataset compared to the reduced datasets. Dice scores were also highest with full MP-MRI for the LGG patients, but statistical significance could not be shown against most reduced MP-MRI datasets. This suggests that inclusion of additional MRI modalities is mainly beneficial for HGG segmentation, as HGGs present with a higher degree of intra-tumoral heterogeneity and more pathological subcompartments. Further research throughout the thesis has therefore mainly focussed on HGG segmentation.

cMRI was found to be the most important MRI modality for brain tumor segmentation: when omitting cMRI from the MP-MRI dataset, the decrease in mean Dice score was the largest for each tissue class compared to the other reduced MP-MRI datasets and statistical significance of the decrease could be shown for all 3 tissue classes (i.e. active tumor, the tumor core and the whole tumor region). This is not surprising, as clinical guidelines for tumor delineation such as the RANO criteria rely upon cMRI data [272].

Statistically significantly lower Dice scores were also found for all 3 tissue classes when omitting MRSI from the MP-MRI feature set. Although MRSI is said to be non-specific in differentiating brain tumors [30], it was found that several MRSI metabolites were among the most specific MRI features to differentiate the pathological tumor compartment from the healthy brain structures. MRSI suffers from some important drawbacks which have till now limited its application in clinical practice. The spatial resolution of brain in-vivo MRSI is poor compared to other MRI modalities, with acquisition voxel size often in the range of $1cm^3$. Such low resolution does not favour an accurate delineation of the tumor volume nor of its subcompartments. Secondly, the MRSI ROI is usually restricted to a rectangular grid within the skull, to avoid

lipid contamination and to keep acquisition time reasonably low. This limitation of spectral volume size often precludes complete coverage of the tumor, which is absolutely mandatory in clinical treatment planning and patient follow-up. To allow full brain analyses and segmentation of the entire tumor volume, MRSI was omitted from the MP-MRI feature set in Chapter 7. There are several promising perspectives that might promote the clinical application of MRSI. An echo-planar MRSI sequence has been presented that does allow full coverage of the whole brain [164]. This technique, which is gradually becoming more widespread in the research field, does allow full coverage of the tumor volume. The spatial resolution of MRSI might be improved with the advent of higher magnetic field strength (7T) scanners. However, to improve performance at higher magnetic field strength, techniques have to be developed to deal with issues like increased magnetic susceptibility effects and shortening of T_2- and T_2^*-relaxation times [9]. Alternatively, more accurate spatial resolution might be obtained by upsampling through super-resolution methods, for instance by incorporating imaging information from other MRI modalities [110].

By including PWI into the MP-MRI feature set, Dice scores only increased significantly for the whole tumor region. Although HGGs often show elevated micro-vascularization in the viable tumor region, rCBV was found to be one of the most heterogeneous MRI features in the active tumor region. While HGGs may focally show an extreme angiogenic response, quantitative studies revealed that the micro-vascular density in many regions of viable tumor is in the range of that of normal gray or white matter [52]. Such lowly perfused tumor areas were not always delineated as tumor when including PWI, leading to lower Dice scores for active tumor and the tumor core. It is possible, however, that the hNMF methodology did not fully exploit the potential of PWI for enhancing tumor segmentation. So far, hNMF has considered one source vector to model each tissue class. Such a model might not work well when the true decision boundaries between the tissue classes are non-linear. It is common practice for instance in brain tumor segmentation to treat enhancing and non-enhancing tumor (based on T1c) as separate tissue classes. Similarly, it might be useful to provide 2 source vectors for the active tumor region when only part of the tumor shows highly elevated rCBV values. This more advanced kind of tissue modeling was already considered with the semi-automated NMF-based segmentation framework in Chapter 7, where several voxels from the same tissue region can be selected and will initiate multiple source vectors if their feature vectors show clear differences.

By including the extensive DKI protocol into the UZ Leuven MP-MRI dataset, Dice scores increased significantly for the tumor core and for the whole tumor region. In general, DWI is mainly helpful in differentiating edema, which has been proven challenging throughout this thesis: edema was regularly confounded

with non-enhancing tumor, gray matter, or even necrosis in the absence of MRSI data. FA and MK are helpful in differentiating white matter tracts that are displaced or distorted but intact from those that are infiltrated or destroyed by neighboring tumor. FA also adds specificity to the detection of the tumor core, as the local FA signal is lost as a result of the destruction of the neuronal structure. An important consideration is the long acquisition time of the DKI protocol: the DKI sequence took 18 minutes, out of a total of 45 minutes for the entire UZ Leuven scanning protocol. A more basic and much shorter DTI sequence would also allow the quantification of MD and FA, thereby only losing the MK parameter in the MP-MRI feature set. Although the added value of MK on the segmentation results has not been explicitly reported in this thesis, experience has learnt that the impact of MK alone is rather low. Although other aspects of the clinical management of gliomas must be considered as well, in terms of segmentation performance it might be preferable to only include a DTI sequence, thereby considerably reducing the acquisition time.

As NMF poses a non-convex optimization problem, the obtained factorization will be a local rather than a global minimum of the objective function. The final result therefore depends on the initialization of the factor matrices. Random initialization is the benchmark and is used in many NMF studies, but it does not warrant robust and reproducible results. In Chapter 4, repetitive random initialization was used for hNMF and the final result was selected based on a sparsity criterion. This approach of course increased computation time dramatically. Alternatively, more advanced initialization strategies might be considered. In **Chapter 5**, the successive projection algorithm (SPA) was proposed as an initialization method for (h)NMF. SPA builds on convex geometry and allocates endmembers based on successive orthogonal subspace projections of the input data. Whereas SPA has commonly been applied as a direct endmember extraction tool, improved segmentation results were found when using SPA as an initialization method, as it allows further enhancement of the sources during the NMF iterative procedure. Comparison with common initialization methods showed that SPA achieves similar segmentation quality and it is competitive in terms of convergence rate.

Computational efficiency of the hNMF algorithm was further improved in **Chapter 6**. The iterative thresholding procedure used for subdividing the voxels over both sources after the first level of NMF, was replaced by a k-means clustering step. By doing so, we were able to make the hNMF computation approximately 10 times faster while maintaining segmentation performance. This has brought hNMF computation time down to the range of other common unsupervised classifiers. In terms of segmentation quality, hNMF was also compared against several unsupervised classifiers, including 2 single-level NMF methods, fuzzy c-means clustering, Gaussian mixture modeling and spectral

clustering. All methods were applied to the MP-MRI data of the UZ Gent and UZ Leuven HGG patients. hNMF globally performed best, with the highest mean Dice score in 17 out of 20 comparisons (2 datasets, 5 tissue classes, 2 initialization methds). hNMF also provided robust segmentation, with relatively low standard deviations and the lowest number of undetected cases for nearly all tissue classes. The hierarchical structure of hNMF gives it some advantages over the other unsupervised methods. After the first level of NMF, tissue classes that are most similar in terms of MP-MRI feature vectors are being assigned to the same source. At the second level, hNMF can then focus on the more subtle differences between the similar tissue classes. This property has allowed hNMF for instance to better separate edema from non-enhancing tumor, which is known to be a challenging segmentation problem. hNMF was also found to better differentiate small tissue regions in some cases. hNMF is probably better able to deal with data imbalance than single-level unsupervised methods, as data imbalance is reduced after the input data have been subdivided into 2 groups after the first level of NMF.

Unsupervised brain tumor segmentation requires the incorporation of prior knowledge into the algorithm, to be able to obtain competitive performance as well as for automation. However, incorporation of prior knowledge often limits the segmentation algorithm to a specific subset of gliomas or to a specific set of MRI features. In **Chapter 7**, a semi-automated framework for brain tumor segmentation was presented based on regularized NMF. L1-regularization was incorporated into the NMF objective function to promote spatial consistency and sparseness of the tissue abundance maps. The pathological sources were initialized through user-defined voxel selection. Knowledge about the spatial location of the selected voxels was combined with tissue adjacency constraints in a post-processing step to enhance the segmentation quality. The incorporated prior knowledge made minimal assumptions regarding tumor appearance, size, shape or location. As a result, the proposed semi-automated segmentation framework is applicable to any type of glioma and to any MP-MRI dataset. Competitive segmentation results were reported on full imaging slices of the UZ Gent dataset, including cMRI, PWI and DWI data, but also on the cMRI data alone. One of the main limitations of semi-automated NMF-based segmentation, is that is does not provide reproducible results. Therefore, robustness against user input variability is a key aspect of the algorithm. User input variability was simulated by repeatedly selecting random voxels from the tumor subcompartments, based on manual segmentation by a radiologist. Robust segmentation results were found for most patients, although low outlier Dice scores were found in several cases. It was concluded that careful voxel selection is mandatory to avoid sub-optimal segmentation.

In **Chapter 8**, the semi-automated segmentation method using L1-regularized

NMF was applied to the publicly available BRATS 2013 Leaderboard dataset. This has allowed a direct comparison to state-of-the art segmentation methods. Our semi-automated NMF-based approach outperformed all other methods in terms of mean Dice score for the enhancing tumor region. Competitive results were also found for the tumor core, but relatively low Dice scores were found for the whole tumor region. One of the main issues that still need to be addressed, is the relatively high amount of false positives in gray matter areas in some patients, which were falsely labelled as edema. Compared to Chapter 7, a more realistic attempt was made to assess intra-user variability of the segmentation results by repeating the analyses several times using different voxel selection by the same radiologist. Although Dice score variability was below 10% in 75% of the cases and below 15% in 86% of the cases, some patients still showed considerable differences among the different runs, with changes in Dice score close to 20% or higher. The vast majority of segmentation algorithms that participated in the BRATS 2012 and 2013 challenges are based on supervised classifiers. Nevertheless, we have shown, based on our results and other relevant work from literature, that unsupervised segmentation methods can be highly competitive when they incorporate proper prior knowledge.

9.2 Future perspectives

The ultimate goal of this work is to contribute to a clinical tool for (semi-)automated brain tumor segmentation. Mainly 2 NMF-based segmentation methods have been developed throughout this thesis: hNMF and semi-automated L1-regularized NMF. hNMF was not considered in the semi-automated segmentation framework for 2 reasons: firstly, hNMF does not allow inclusion of spatial regularization, as the spatial relationship between the voxels is lost after subdividing the voxels prior to the second level of NMF. Secondly, although the total number of tissue types is known after voxel selection by the user, rank estimation would not be fully automated with hNMF. The hNMF algorithm currently requires the user to manually specify the number of tissue types represented by each source after the first level of NMF, based upon visual inspection of the abundance maps in combination with the cMRI data. Automatic rank estimation poses a challenging problem, and to date none of the proposed methods are full-proof [267]. In the current work, the rank estimation problem is further complicated by the large discrepancy in spatial resolution among the MRI modalities, resulting in varying degrees of partial volume effects. Although not reported in this thesis, several common methods for **automatic rank estimation** were explored, but none of them were successful: principal component analysis, Bayesian information criteria

[7], Fogel and Young's volume-based method [74] and automatic relevance determination [282].

Somehow, the rank estimation problem needs to be solved for hNMF to become an automated and reproducible segmentation method. However, one has to wonder whether the 'correct' rank is necessarily equal to the number of tissue types present within the ROI. As mentioned before, segmentation might be more successful by allowing several sources to model the same tissue class: for instance, active tumor might consist of both enhancing and non-enhancing tumor, and some regions of the tumor might show hyper-perfusion whereas other regions don't exhibit vascular ingrowth. Voxels near the boundary of a tissue region will show different behaviour in terms of texture features, when these are included into the feature set. From this point of view, the number of sources not only depends on the number of tissue classes within the ROI, but also on the amount of within-tissue variability. In the semi-automated NMF-based framework, for instance, the number of sources used to model each tumor subcompartment depends on the correlation between the feature vectors of the selected voxels. Another straightforward way of dealing with rank estimation, is to use an **over-determined model** with a fixed number of sources. The main challenge is then to automatically fuse the sources belonging to the same tissue class, based on prior knowledge [104, 115]. The hNMF model in its current form still struggles with data imbalance, especially when applied to full 3D imaging datasets. When the number of sources is restricted to the number of tissue types, small pathological tissue regions won't be exclusively modeled by their own source vector. Using an over-determined model might at the same time solve this problem as well. In order for the source vectors to remain physically meaningful and still correspond to tissue-specific signatures with such an over-determined model, one would probably have to use an NMF implementation that promotes sparseness at the second level of hNMF, such as e.g. Convex NMF. However, maintaining computational efficiency of the hNMF algorithm is another crucial aspect that needs to be taken into account throughout its further development.

If one wants to build a fully automated unsupervised segmentation method with no user interaction, for instance based on hNMF, a popular approach nowadays is to incorporate prior knowledge in the form of a **healthy brain atlas**. Several studies have reported excellent results based on an unsupervised classifier in combination with an atlas [115, 129]. An important drawback of such an approach, however, is that non-rigid co-registration to the atlas requires long computation times. Alternatively, one could incorporate **prior knowledge about the appearance of certain tissue types in specific MR images**. NMF-based segmentation methods that have been developed throughout this thesis have been kept as flexible as possible, allowing them to cope with any

MP-MRI dataset. This philosophy, however, has severely restricted the amount of prior knowledge that could be applied, as no assumptions could be made about the availability of any type of MR images. It seems reasonable to assume that cMRI data such as T1c and FLAIR will virtually always be available, as they are crucial for obtaining good segmentation results. One might then for instance apply a threshold to the FLAIR images, to get a rough initial estimate of the entire pathological region as it is known to be hyperintense on FLAIR [115]. Making assumptions about the availability of certain MR images to incorporate related prior knowledge is probably mandatory if one wants to avoid using a brain atlas or user input.

Competitive segmentation results have been reported for semi-automated L1-regularized NMF, both on the UZ Gent as well as on the BRATS Leaderboard dataset. Nevertheless, there is still room for improvement, in particular in segmenting the whole tumor region. One of the remaining issues is the relatively high amount of false positives for edema. These mainly consist of gray matter areas, which could not be separated from true edema by morphological post-processing when connected to the edema region. One possible way of removing these false positives, is to apply **more stringent connectivity constraints**. Weak connectivity might for instance be removed by applying morphological erosion to the segmentation masks. Alternatively, incorporation of **additional prior knowledge** might also further improve segmentation performance. Specification of the maximal tumor diameter by the user, for instance, is also considered in several segmentation algorithms [96, 129] and might further help in the removal of false positives. Another possibility, is to also ask the user to select voxels in the healthy brain compartments. This would allow for instance to explicitly initialize one or more gray matter sources, which could also further reduce the amount of false positives. The main disadvantage of initializing all the sources based on user input, is that data variability might not always be fully covered by the user. For instance, larger artefactial regions have to be modeled as well, to avoid false positives. Inclusion of PWI requires blood vessels to be modeled as a separate tissue class, whereas blood vessels would be included in the CSF tissue class when only cMRI is considered.

Based on the reported PPV and sensitivity values on the BRATS Leaderboard dataset in Chapter 8, it was found that semi-automated NMF has a relatively high amount of false negatives compared to other segmentation methods. This is probably related to the fact that in most cases we model each tissue class by only one source vector, and the NMF model does not incorporate within-class variability. The semi-automated NMF approach does allow for **several sources per tissue class**, but an arbitrary threshold is applied to the correlation coefficient between candidate source vectors to decide whether or not they should be merged. A more **advanced and adaptive method of defining**

the threshold for merging the sources should therefore be sought, also taking into account correlation values with respect to the other tissue classes. Alternatively, one might consider an NMF-like model which does also allow for modeling within-tissue variability, such as finite mixture models.

Besides segmentation performance, an important concern for semi-automated NMF is the **robustness against user input variability**. It is possible that improvement of the algorithm in terms of performance also leads to a lower variability in the Dice scores. Furthermore, the influence of the number of selected voxels on the segmentation result has not yet been investigated. A relatively low amount of voxels per tissue class, usually between 3 and 5, has been considered in Chapters 7 and 8. It is possible that robustness of the results increases by selecting more voxels. Due to the merging of the candidate source vectors based on their correlation, complexity of the NMF model probably only mildly increases with the number of selected voxels.

Another big challenge that is remaining for semi-automated NMF is the **reduction of computation time**. Whereas immediate feedback is not mandatory for treating brain tumors, segmentation algorithms taking one hour or more per patient might not be accepted in clinical practice. Using semi-automated L1-regularized NMF, the average computation time per patient on the BRATS dataset was 75 minutes. The NMF computation itself takes up most of the time. **More efficient NMF algorithms** do exist, but SDF NMF does have several benefits: it is sensitive to the initialization, which is an asset in this case, and it allows straightforward incorporation of (L1-)regularization. A more efficient NMF implementation might have to be sought, with minimal or no loss of segmentation performance. It must be noted that inclusion of the L1-regularization term has approximately doubled the computation time. Alternatively, a **more efficient algorithm for promoting spatial consistency** such as MRF might be considered in a post-processing step. Another option would be to impose less stringent convergence criteria, on the condition that this does not compromise segmentation performance. In-plane spatial subsampling by a factor 2 was already implemented, and was found not to affect segmentation performance. As there is a lot of redundancy in the MR images, a **more radical subsampling strategy** might be considered as well. However, such an approach would preclude the inclusion of spatial regularization, as the spatial structure of the voxels would be lost. Again, spatial regularization might still be considered in a post-processing step.

Whereas all datasets that were analysed throughout this thesis consisted of untreated glioma images, **post-treatment monitoring** of volumetric changes of the tumor and its subcompartments is also crucial, in order to make accurate prognosis and assess success of therapy at the earliest possible stage. Therefore, automated routines should also be evaluated in longitudinal settings, where the

change of tumor structures over time is of primary relevance. In the BRATS 2016 challenge, for the first time longitudinal imaging data with multiple observations per patient are included into the test. Post-treatment segmentation will pose additional challenges, such as the distinction of pseudo-progression, radiation necrosis, and possibly pseudo-response as a result of anti-angiogenic therapy. The added value of individual MRI modalities for tumor segmentation will probably be different compared to the pre-treatment stage and will have to be re-evaluated. Furthermore, segmentation algorithms might benefit from **exploiting the longitudinal information**, as successive delineations are expected to be similar. It is therefore probable that segmentation algorithms that perform very well on pre-treatment images might have to be adapted for optimal performance in longitudinal settings. Future work should therefore focus on the performance of NMF-based segmentation methods on longitudinal data, and explore whether NMF-based segmentation might benefit from adaptations to the algorithm by including temporal information.

NMF provides the relative abundance of the different tissue classes in each voxel. In this thesis, NMF results were not fully exploited, as the abundance maps were converted into absolute tissue segmentation, assigning each voxel to only one tissue class. Nevertheless, NMF abundance maps might also directly be useful in the clinical management of gliomas. The highest abundances of active tumor might serve well in asessing the most malignant part of a tumor for **biopsy targeting**. The active tumor map might also allow the detection of areas of **peri-tumoral infiltration**. Relatively low but non-zero abundances of active tumor might indicate tumor infiltration. Probably, a subset of MP-MRI features would have to be considered, as not all features are of added value in assessing tumor malignancy or infiltration. The potential of NMF regarding biopsy targeting and detecting infiltration has not been explored because of a lack of proper means of validation. Assessment of the most malignant part of a tumor and of tumor infiltration can currently not be made based on imaging data alone. It would require either pre-clinical studies or neuro-navigated biopsy sampling in combination with advanced imaging protocols to draw firm conclusions about these issues.

Appendix A

Validation scores individual UZ Leuven patients

Dice scores and correlation coefficients averaged over the UZ Leuven patient cohort were reported for hNMF in Chapter 4. This appendix contains the validation scores of the individual patients, supporting the results from Chapter 4. Table A.1 reports the Dice scores for the HGG patients for the full MP-MRI dataset, as well as for the reduced MRI datasets. Table A.2 reports the correlation coefficients for the HGG patients for all considered MRI datasets. Table A.3 reports the Dice scores for the LGG patients for all considered MRI datasets. Table A.4 reports the correlation coefficients for the LGG patients for all considered MRI datasets.

	Dice score active tumor [%]						Dice score tumor core [%]						Dice score whole tumor [%]					
	All	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**	**All**	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**	**All**	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**
PAT01	74	76	76	56	66	51	74	76	76	56	66	51	74	76	76	56	66	51
PAT02	93	91	93	84	93	91	93	91	93	84	93	91	93	91	93	84	93	91
PAT03	88	70	91	84	87	89	88	70	91	84	87	89	88	70	91	84	87	89
PAT04	81	69	83	85	66	82	98	89	98	92	98	95	90	79	90	86	89	76
PAT05	83	64	86	81	48	84	89	81	90	88	51	76	80	67	81	77	43	69
PAT06	81	67	87	78	78	77	92	89	95	88	86	58	83	80	83	84	84	79
PAT07	68	39	78	74	70	52	76	70	74	65	74	48	98	98	98	98	97	97
PAT08	86	67	70	89	86	76	86	67	70	89	86	76	59	50	52	61	58	46
PAT09	82	50	86	77	90	87	93	80	95	90	96	91	93	86	92	88	92	47
PAT10	70	66	58	76	51	55	89	67	52	71	74	74	93	77	75	71	85	56
PAT11	59	70	76	58	68	49	79	92	85	76	84	72	69	76	63	70	68	66
PAT12	75	55	68	70	66	60	89	77	84	80	82	62	89	77	84	80	82	62
PAT13	64	44	20	58	43	38	54	40	36	48	40	39	54	40	36	48	40	39
PAT14	88	84	88	92	85	73	88	84	88	92	85	73	92	88	88	91	86	79
Mean	**78**	**65**	**76**	**76**	**71**	**69**	**85**	**77**	**81**	**79**	**79**	**71**	**83**	**75**	**79**	**77**	**76**	**68**

Table A.1: Dice scores for active tumor, the tumor core and the whole tumor region for the UZ Leuven HGG patients. Results are reported for the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only.

	Correlation coefficient tumor						Correlation coefficient necrosis						Correlation coefficient edema					
	All	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**	**All**	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**	**All**	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**
PAT01	0.99	0.99	0.99	0.99	0.96	0.82	-	-	-	-	-	-	-	-	-	-	-	-
PAT02	0.99	1.00	1.00	1.00	1.00	1.00	-	-	-	-	-	-	-	-	-	-	-	-
PAT03	0.99	0.97	0.99	0.98	0.98	0.99	-	-	-	-	-	-	-	-	-	-	-	-
PAT04	0.71	0.89	0.82	0.95	0.71	0.72	0.96	0.96	0.95	0.97	0.97	0.70	0.93	0.98	0.95	0.96	0.77	0.97
PAT05	0.97	0.89	0.99	0.98	0.54	0.81	0.99	0.58	0.90	0.97	0.81	0.93	-	-	-	-	-	-
PAT06	0.78	0.85	0.76	0.93	0.71	0.99	0.98	0.96	0.98	0.97	0.98	0.97	0.97	0.92	0.94	0.99	0.95	-
PAT07	0.91	0.87	0.97	0.94	0.83	0.99	0.98	0.97	0.97	0.86	0.97	1.00	1.00	0.99	1.00	0.99	1.00	0.76
PAT08	0.95	0.97	0.96	0.98	0.96	0.70	-	-	-	-	-	-	0.92	0.84	0.91	0.99	0.92	0.99
PAT09	0.80	0.88	0.68	0.98	0.95	0.60	0.99	1.00	0.99	1.00	0.99	0.96	0.98	0.89	0.98	0.96	1.00	-
PAT10	0.86	0.76	0.63	0.93	0.67	0.59	0.99	0.97	0.98	0.98	0.99	0.99	0.98	-	-	0.97	-	-
PAT11	0.99	0.96	0.99	0.93	0.99	0.99	0.99	0.98	0.99	0.98	0.98	0.99	0.90	0.92	0.80	0.89	0.92	-
PAT12	0.99	0.98	0.99	0.99	0.71	0.96	0.97	0.95	0.98	0.98	0.96	0.99	-	-	-	-	-	-
PAT13	0.87	0.93	0.77	0.94	0.96	0.99	0.90	0.91	0.98	0.87	0.93	0.91	-	-	-	-	-	-
PAT14	0.98	0.96	0.98	0.98	0.96	0.99	-	-	-	-	-	-	0.99	0.98	0.99	0.99	0.98	0.98
Mean	**0.91**	**0.92**	**0.89**	**0.96**	**0.85**	**0.87**	**0.97**	**0.92**	**0.97**	**0.95**	**0.95**	**0.94**	**0.96**	**0.93**	**0.94**	**0.97**	**0.93**	**0.93**

Table A.2: Correlation coefficients for active tumor, necrosis and edema for the UZ Leuven HGG patients. Results are reported for the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only.

	Dice score whole tumor [%]					
	All	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**
PAT15	89	79	90	90	89	88
PAT16	50	42	52	46	41	38
PAT17	85	61	87	87	87	88
PAT18	76	70	67	84	75	71
PAT19	94	95	94	91	95	92
PAT20	93	87	88	94	95	96
PAT21	91	45	68	89	92	75
PAT22	96	94	96	95	96	96
PAT23	95	95	95	96	96	95
PAT24	76	60	70	68	77	72
Mean	**85**	**73**	**81**	**84**	**84**	**81**

Table A.3: Dice scores for the whole tumor region of the UZ Leuven LGG patients. Results are reported for the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only.

	Correlation coefficient active tumor					
	All	**no-cMRI**	**no-PWI**	**no-DKI**	**no-MRSI**	**cMRI only**
PAT15	0.96	0.95	1.00	0.99	0.97	0.97
PAT16	0.75	0.57	0.72	0.94	0.94	0.97
PAT17	0.88	0.80	0.97	0.98	0.97	0.94
PAT18	0.97	0.97	0.97	0.98	0.99	0.99
PAT19	1.00	1.00	1.00	1.00	0.99	0.99
PAT20	1.00	0.99	1.00	1.00	1.00	0.99
PAT21	0.99	0.73	0.99	0.97	0.99	0.99
PAT22	1.00	1.00	1.00	1.00	1.00	1.00
PAT23	1.00	1.00	1.00	1.00	1.00	1.00
PAT24	0.93	0.97	0.93	0.97	0.93	0.89
Mean	**0.95**	**0.90**	**0.96**	**0.98**	**0.98**	**0.97**

Table A.4: Correlation coefficients for the active tumor region of the UZ Leuven LGG patients. Results are reported for the full MP-MRI dataset, MP-MRI without cMRI, PWI, DKI and MRSI, and for cMRI data only.

Bibliography

[1] Andersson, J. L., and Sotiropoulos, S. N. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage 125* (2016), 1063–1078.

[2] Araújo, M. C. U., Saldanha, T. C. B., Galvão, R. K. H., Yoneyama, T., Chame, H. C., and Visani, V. The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometrics and Intelligent Laboratory Systems 57*, 2 (2001), 65–73.

[3] Aronen, H. J., Gazit, I. E., Louis, D. N., Buchbinder, B. R., Pardo, F. S., Weisskoff, R. M., Harsh, G. R., Cosgrove, G., Halpern, E. F., and Hochberg, F. H. Cerebral blood volume maps of gliomas: comparison with tumor grade and histologic findings. *Radiology 191*, 1 (1994), 41–51.

[4] Arthur, D., and Vassilvitskii, S. k-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (2007), Society for Industrial and Applied Mathematics, pp. 1027–1035.

[5] Assaf, Y., Ben-Bashat, D., Chapman, J., Peled, S., Biton, I., Kafri, M., Segev, Y., Hendler, T., Korczyn, A., Graif, M., et al. High b-value q-space analyzed diffusion-weighted MRI: Application to multiple sclerosis. *Magnetic Resonance in Medicine 47*, 1 (2002), 115–126.

[6] Astrakas, L. G., Zurakowski, D., Tzika, A. A., Zarifi, M. K., Anthony, D. C., De Girolami, U., Tarbell, N. J., and Black, P. M. Noninvasive magnetic resonance spectroscopic imaging biomarkers to predict the clinical grade of pediatric brain tumors. *Clinical Cancer Research 10*, 24 (2004), 8220–8228.

[7] Bai, J., and Ng, S. Determining the number of factors in approximate factor models. *Econometrica 70*, 1 (2002), 191–221.

[8] Barajas Jr, R. F., Chang, J. S., Segal, M. R., Parsa, A. T., McDermott, M. W., Berger, M. S., and Cha, S. Differentiation of recurrent glioblastoma multiforme from radiation necrosis after external beam radiation therapy with dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging 1. *Radiology 253*, 2 (2009), 486–496.

[9] Barker, P. B. Ultra-high field MRSI (7T and beyond). In *Functional Brain Tumor Imaging*. Springer, 2014, pp. 195–209.

[10] Barker, P. B., Breiter, S. N., Soher, B. J., Chatham, J. C., Forder, J. R., Samphilipo, M. A., Magee, C. A., and Anderson, J. H. Quantitative proton spectroscopy of canine brain: in vivo and in vitro correlations. *Magnetic Resonance in Medicine 32*, 2 (1994), 157–163.

[11] Basser, P. J., Mattiello, J., and LeBihan, D. Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance, Series B 103*, 3 (1994), 247–254.

[12] Basser, P. J., Mattiello, J., and LeBihan, D. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal 66*, 1 (1994), 259.

[13] Bauer, S., Fejes, T., Slotboom, J., Wiest, R., Nolte, L.-P., and Reyes, M. Segmentation of brain tumor images based on integrated hierarchical classification and regularization. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2012), Springer, pp. 10–13.

[14] Bauer, S., Wiest, R., Nolte, L.-P., and Reyes, M. A survey of MRI-based medical image analysis for brain tumor studies. *Physics in Medicine and Biology 58*, 13 (2013), R97.

[15] Beck, T. J., Kreth, F. W., Beyer, W., Mehrkens, J. H., Obermeier, A., Stepp, H., Stummer, W., and Baumgartner, R. Interstitial photodynamic therapy of nonresectable malignant glioma recurrences using 5-aminolevulinic acid induced protoporphyrin ix. *Lasers in Surgery and Medicine 39*, 5 (2007), 386–393.

[16] Bell, H., Whittle, I., Walker, M., Leaver, H., and Wharton, S. The development of necrosis and apoptosis in glioma: experimental findings using spheroid culture systems. *Neuropathology and Applied Neurobiology 27*, 4 (2001), 291–304.

[17] Blana, A., Brown, S. C., Chaussy, C., Conti, G. N., Eastham, J. A., Ganzer, R., Murat, F. J., Pasticier, G., Rebillard, X., Rewcastle, J. C., et al. High-intensity focused ultrasound for prostate cancer: comparative definitions of biochemical failure. *BJU International 104*, 8 (2009), 1058–1062.

[18] Bottomley, P. A. Spatial localization in NMR spectroscopy in vivo. *Annals of the New York Academy of Sciences 508*, 1 (1987), 333–348.

[19] Boutsidis, C., and Gallopoulos, E. SVD based initialization: A head start for nonnegative matrix factorization. *Pattern Recognition 41*, 4 (2008), 1350–1362.

[20] Boxerman, J., Schmainda, K., and Weisskoff, R. Relative cerebral blood volume maps corrected for contrast agent extravasation significantly correlate with glioma tumor grade, whereas uncorrected maps do not. *American Journal of Neuroradiology 27*, 4 (2006), 859–867.

[21] Boxerman, J. L., Hamberg, L. M., Rosen, B. R., and Weisskoff, R. M. MR contrast due to intravascular magnetic susceptibility perturbations. *Magnetic Resonance in Medicine 34*, 4 (1995), 555–566.

[22] Brandão, L. A., Shiroishi, M. S., and Law, M. Brain tumors: a multimodality approach with diffusion-weighted imaging, diffusion tensor imaging, magnetic resonance spectroscopy, dynamic susceptibility contrast and dynamic contrast-enhanced magnetic resonance imaging. *Magnetic Resonance Imaging Clinics of North America 21*, 2 (2013), 199–239.

[23] Brandes, A. A., Franceschi, E., Tosoni, A., Blatt, V., Pession, A., Tallini, G., Bertorelle, R., Bartolini, S., Calbucci, F., Andreoli, A., et al. MGMT promoter methylation status can predict the incidence and outcome of pseudoprogression after concomitant radiochemotherapy in newly diagnosed glioblastoma patients. *Journal of Clinical Oncology 26*, 13 (2008), 2192–2197.

[24] Brandes, A. A., Tosoni, A., Spagnolli, F., Frezza, G., Leonardi, M., Calbucci, F., and Franceschi, E. Disease progression or pseudoprogression after concomitant radiochemotherapy treatment: pitfalls in neurooncology. *Neuro-Oncology 10*, 3 (2008), 361–367.

[25] Brandsma, D., and van den Bent, M. J. Pseudoprogression and pseudoresponse in the treatment of gliomas. *Current Opinion in Neurology 22*, 6 (2009), 633–638.

[26] Brito, M., Chavez, E., Quiroz, A., and Yukich, J. Connectivity of the mutual k-nearest-neighbor graph in clustering and outlier detection. *Statistics & Probability Letters 35*, 1 (1997), 33–42.

[27] Bro, R., and De Jong, S. A fast non-negativity-constrained least squares algorithm. *Journal of Chemometrics 11*, 5 (1997), 393–401.

[28] Bulakbasi, N., Guvenc, I., Onguru, O., Erdogan, E., Tayfun, C., and Ucoz, T. The added value of the apparent diffusion coefficient calculation to magnetic resonance imaging in the differentiation and grading of malignant brain tumors. *Journal of Computer Assisted Tomography 28*, 6 (2004), 735–746.

[29] Bulakbasi, N., Kocaoglu, M., Farzaliyev, A., Tayfun, C., Ucoz, T., and Somuncu, I. Assessment of diagnostic accuracy of perfusion MR imaging in primary and metastatic solitary malignant brain tumors. *American Journal of Neuroradiology 26*, 9 (2005), 2187–2199.

[30] Bulik, M., Jancalek, R., Vanicek, J., Skoch, A., and Mechl, M. Potential of MR spectroscopy for assessment of glioma grading. *Clinical Neurology and Neurosurgery 115*, 2 (2013), 146–153.

[31] Bulik, M., Kazda, T., Slampa, P., and Jancalek, R. The diagnostic ability of follow-up imaging biomarkers after treatment of glioblastoma in the temozolomide era: implications from proton MR spectroscopy and apparent diffusion coefficient mapping. *BioMed Research International 2015* (2015).

[32] Byrnes, T. J., Barrick, T. R., Bell, B. A., and Clark, C. A. Diffusion tensor imaging discriminates between glioblastoma and cerebral metastases in vivo. *NMR in Biomedicine 24*, 1 (2011), 54–60.

[33] Cairncross, G., Berkey, B., Shaw, E., Jenkins, R., Scheithauer, B., Brachman, D., Buckner, J., Fink, K., Souhami, L., Laperierre, N., et al. Phase iii trial of chemotherapy plus radiotherapy compared with radiotherapy alone for pure and mixed anaplastic oligodendroglioma: Intergroup radiation therapy oncology group trial 9402. *Journal of Clinical Oncology 24*, 18 (2006), 2707–2714.

[34] Calamante, F. Quantification of dynamic susceptibility contrast T2* MRI in oncology. In *Dynamic contrast-enhanced magnetic resonance imaging in oncology*. Springer, 2005, pp. 53–67.

[35] Calamante, F., Gadian, D. G., and Connelly, A. Delay and dispersion effects in dynamic susceptibility contrast MRI: simulations using singular value decomposition. *Magnetic Resonance in Medicine 44*, 3 (2000), 466–473.

[36] Calamante, F., Thomas, D. L., Pell, G. S., Wiersma, J., and Turner, R. Measuring cerebral blood flow using magnetic resonance

imaging techniques. *Journal of Cerebral Blood Flow & Metabolism 19*, 7 (1999), 701–735.

[37] Cao, Y., Tsien, C. I., Nagesh, V., Junck, L., Ten Haken, R., Ross, B. D., Chenevert, T. L., and Lawrence, T. S. Clinical investigation survival prediction in high-grade gliomas by MRI perfusion before and during early stage of RT. *International Journal of Radiation Oncology* Biology* Physics 64*, 3 (2006), 876–885.

[38] Casalino, G., Del Buono, N., and Mencar, C. Subtractive clustering for seeding non-negative matrix factorizations. *Information Sciences 257* (2014), 369–387.

[39] Cha, S. Update on brain tumor imaging: from anatomy to physiology. *American Journal of Neuroradiology 27*, 3 (2006), 475–487.

[40] Cha, S., Knopp, E. A., Johnson, G., Wetzel, S. G., Litt, A. W., and Zagzag, D. Intracranial mass lesions: Dynamic contrast-enhanced susceptibility-weighted echo-planar perfusion MR imaging 1. *Radiology 223*, 1 (2002), 11–29.

[41] Cha, S., Tihan, T., Crawford, F., Fischbein, N. J., Chang, S., Bollen, A., Nelson, S. J., Prados, M., Berger, M. S., and Dillon, W. P. Differentiation of low-grade oligodendrogliomas from low-grade astrocytomas by using quantitative blood-volume measurements derived from dynamic susceptibility contrast-enhanced MR imaging. *American Journal of Neuroradiology 26*, 2 (2005), 266–273.

[42] Chan, J. L., Lee, S. W., Fraass, B. A., Normolle, D. P., Greenberg, H. S., Junck, L. R., Gebarski, S. S., and Sandler, H. M. Survival and failure patterns of high-grade gliomas after three-dimensional conformal radiotherapy. *Journal of Clinical Oncology 20*, 6 (2002), 1635–1642.

[43] Chan, Y.-l., Yeung, D. K., Leung, S.-f., and Chan, P.-n. Diffusion-weighted magnetic resonance imaging in radiation-induced cerebral necrosis: apparent diffusion coefficient in lesion components. *Journal of Computer Assisted Tomography 27*, 5 (2003), 674–680.

[44] Chard, D., Griffin, C., McLean, M., Kapeller, P., Kapoor, Rand Thompson, A., and Miller, D. Brain metabolite changes in cortical grey and normal-appearing white matter in clinically early relapsing-remitting multiple sclerosis. *Brain 125*, 10 (2002).

[45] Cheng, L. L., Anthony, D. C., Comite, A. R., Black, P. M., Tzika, A. A., and Gonzalez, R. G. Quantification of

microheterogeneity in glioblastoma multiforme with ex vivo high-resolution magic-angle spinning (HRMAS) proton magnetic resonance spectroscopy. *Neuro-Oncology 2*, 2 (2000), 87–95.

[46] Chernov, M. F., Muragaki, Y., Ochiai, T., Taira, T., Ono, Y., Usukura, M., Maruyama, T., Nakaya, K., Nakamura, R., Iseki, H., et al. Spectroscopy-supported frame-based image-guided stereotactic biopsy of parenchymal brain lesions: comparative evaluation of diagnostic yield and diagnostic accuracy. *Clinical Neurology and Neurosurgery 111*, 6 (2009), 527–535.

[47] Cho, K., Wang, X., Nie, S., Shin, D. M., et al. Therapeutic nanoparticles for drug delivery in cancer. *Clinical Cancer Research 14*, 5 (2008), 1310–1316.

[48] Choi, C., Ganji, S. K., DeBerardinis, R. J., Dimitrov, I. E., Pascual, J. M., Bachoo, R., Mickey, B. E., Malloy, C. R., and Maher, E. A. Measurement of glycine in the human brain in vivo by 1H-MRS at 3T: application in brain tumors. *Magnetic Resonance in Medicine 66*, 3 (2011), 609–618.

[49] Chu, M., Diele, F., Plemmons, R., and Ragni, S. Optimality, computation, and interpretation of nonnegative matrix factorizations. In *SIAM Journal on Matrix Analysis* (2004), Citeseer.

[50] Cichocki, A., Zdunek, R., and Amari, S.-i. Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization. In *Independent Component Analysis and Signal Separation*, vol. LCNS 4666. Springer, 2007, pp. 169–176.

[51] Cichocki, A., Zdunek, R., Phan, A. H., and Amari, S.-i. *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation.* John Wiley & Sons, 2009.

[52] Claes, A., Idema, A. J., and Wesseling, P. Diffuse glioma growth: a guerilla war. *Acta Neuropathologica 114*, 5 (2007), 443–458.

[53] Clarke, J. L., and Chang, S. Pseudoprogression and pseudoresponse: challenges in brain tumor imaging. *Current Neurology and Neuroscience Reports 9*, 3 (2009), 241–246.

[54] Clatz, O., Bondiau, P.-Y., Delingette, H., Sermesant, M., Warfield, S. K., Malandain, G., and Ayache, N. *Brain tumor growth simulation.* PhD thesis, INRIA, 2004.

[55] Cocosco, C. A., Kollokian, V., Kwan, R. K.-S., Pike, G. B., and Evans, A. C. Brainweb: Online interface to a 3D MRI simulated brain database. In *NeuroImage* (1997), Citeseer.

[56] Cordier, N., Menze, B., Delingette, H., and Ayache, N. Patch-based segmentation of brain tissues. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2013), Springer, pp. 6–17.

[57] Croitor Sava, A. R., Sima, D. M., Poullet, J.-B., Wright, A. J., Heerschap, A., and Van Huffel, S. Exploiting spatial information to estimate metabolite levels in two-dimensional MRSI of heterogeneous brain lesions. *NMR in Biomedicine 24*, 7 (2011), 824–835.

[58] Daumas-Duport, C., Meder, J., Monsaingeon, V., Missir, O., Aubin, M., and Szikla, G. Cerebral gliomas: malignancy, limits and spatial configuration. comparative data from serial stereotaxic biopsies and computed tomography (a preliminary study based on 50 cases). *Journal of Neuroradiology 10*, 1 (1983), 51.

[59] DeAngelis, L. M. Brain tumors. *New England Journal of Medicine 344*, 2 (2001), 114–123.

[60] Dempster, A. P., Laird, N. M., and Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (methodological)* (1977), 1–38.

[61] Detre, J. A., Rao, H., Wang, D. J., Chen, Y. F., and Wang, Z. Applications of arterial spin labeled MRI in the brain. *Journal of Magnetic Resonance Imaging 35*, 5 (2012), 1026–1037.

[62] Devarajan, K. Nonnegative matrix factorization: an analytical and interpretive tool in computational biology. *PLoS Comput Biol 4*, 7 (2008), e1000029.

[63] Dhermain, F. Radiotherapy of high-grade gliomas: current standards and new concepts, innovations in imaging and radiotherapy, and new therapeutic approaches. *Chin J Cancer 33*, 1 (2014), 16–24.

[64] Di Costanzo, A., Scarabino, T., Trojsi, F., Giannatempo, G. M., Popolizio, T., Catapano, D., Bonavita, S., Maggialetti, N., Tosetti, M., Salvolini, U., d'Angelo, V., and Tedeschi, G. Multiparametric 3T MR approach to the assessment of cerebral gliomas: tumor extent and malignancy. *Neuroradiology 48*, 9 (2006), 622–631.

[65] Diehl, F., Gales, M. J. F., Liu, X., Tomalin, M., Woodland, P. C., et al. Word boundary modelling and full covariance gaussians for arabic speech-to-text systems. In *INTERSPEECH* (2011), pp. 777–780.

[66] Ding, C., Li, T., Jordan, M., et al. Convex and semi-nonnegative matrix factorizations. *IEEE Transactions on Pattern Analysis and Machine Intelligence 32*, 1 (2010), 45–55.

[67] Doyle, S., Vasseur, F., Dojat, M., and Forbes, F. Fully automatic brain tumor segmentation from multiple MR sequences using hidden Markov fields and variational EM. *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2013), 18–22.

[68] Drevelegas, A., and Papanikolaou, N. *Imaging modalities in brain tumors*. Springer, 2011.

[69] Ebel, A., and Maudsley, A. A. Detection and correction of frequency instabilities for volumetric 1H echo-planar spectroscopic imaging. *Magnetic Resonance in Medicine 53*, 2 (2005), 465–469.

[70] Eis, M., Els, T., Hoehn-Berlage, M., and Hossmann, K.-A. Quantitative diffusion MR imaging of cerebral tumor and edema. In *Brain Edema IX*. Springer, 1994, pp. 344–346.

[71] Ellingson, B. M., Cloughesy, T. F., Lai, A., Nghiemphu, P. L., Mischel, P. S., and Pope, W. B. Quantitative volumetric analysis of conventional MRI response in recurrent glioblastoma treated with bevacizumab. *Neuro-Oncology* (2011), noq206.

[72] Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., et al. 3D slicer as an image computing platform for the quantitative imaging network. *Magnetic Resonance Imaging 30*, 9 (2012), 1323–1341.

[73] Fletcher-Heath, L. M., Hall, L. O., Goldgof, D. B., and Murtagh, F. R. Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. *Artificial Intelligence in Medicine 21*, 1 (2001), 43–63.

[74] Fogel, P., Young, S. S., Hawkins, D. M., and Ledirac, N. Inferential, robust non-negative matrix factorization analysis of microarray data. *Bioinformatics 23*, 1 (2007), 44–49.

[75] Forsyth, P. A., and Posner, J. B. Headaches in patients with brain tumors: a study of 111 patients. *Neurology 43*, 9 (1993), 1678–1678.

[76] Fountas, K. N., Kapsalaki, E. Z., Vogel, R. L., Fezoulidis, I., Robinson, J. S., and Gotsis, E. D. Noninvasive histologic grading of solid astrocytomas using proton magnetic resonance spectroscopy. *Stereotactic and Functional Neurosurgery 82*, 2-3 (2004), 90–97.

[77] Friedman, H. S., Prados, M. D., Wen, P. Y., Mikkelsen, T., Schiff, D., Abrey, L. E., Yung, W. A., Paleologos, N., Nicholas, M. K., Jensen, R., et al. Bevacizumab alone and in combination with irinotecan in recurrent glioblastoma. *Journal of Clinical Oncology 27*, 28 (2009), 4733–4740.

[78] Furuya, S., Naruse, S., Ide, M., Morishita, H., Kizu, O., Ueda, S., and Maeda, T. Evaluation of metabolic heterogeneity in brain tumors using 1H-chemical shift imaging method. *NMR in Biomedicine 10*, 1 (1997), 25–30.

[79] Gering, D. T., Grimson, W. E. L., and Kikinis, R. Recognizing deviations from normalcy for brain tumor segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2002), Springer, pp. 388–395.

[80] Gillis, N. Successive nonnegative projection algorithm for robust nonnegative blind source separation. *SIAM Journal on Imaging Sciences 7*, 2 (2014), 1420–1450.

[81] Gillis, N., et al. *Nonnegative matrix factorization: Complexity, algorithms and applications*. PhD thesis, UCL, 2011.

[82] Gillis, N., and Glineur, F. Accelerated multiplicative updates and hierarchical ALS algorithms for nonnegative matrix factorization. *Neural Computation 24*, 4 (2012), 1085–1105.

[83] Gillis, N., Kuang, D., and Park, H. Hierarchical clustering of hyperspectral images using rank-two nonnegative matrix factorization. *IEEE Transactions on Geoscience and Remote Sensing 53*, 4 (2015), 2066–2078.

[84] Gillis, N., Vavasis, S., et al. Fast and robust recursive algorithms for separable nonnegative matrix factorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence 36*, 4 (2014), 698–714.

[85] Ginsberg, L. E., Fuller, G. N., Hashmi, M., Leeds, N. E., and Schomer, D. F. The significance of lack of MR contrast enhancement of supratentorial brain tumors in adults: histopathological evaluation of a series. *Surgical Neurology 49*, 4 (1998), 436–440.

[86] Goebell, E., Fiehler, J., Ding, X.-Q., Paustenbach, S., Nietz, S., Heese, O., Kucinski, T., Hagel, C., Westphal, M., and Zeumer, H. Disarrangement of fiber tracts and decline of neuronal density correlate in glioma patients—a combined diffusion tensor imaging and 1H-MR spectroscopy study. *American Journal of Neuroradiology 27*, 7 (2006), 1426–1431.

[87] Gordillo, N., Montseny, E., and Sobrevilla, P. State of the art survey on MRI brain tumor segmentation. *Magnetic Resonance Imaging 31*, 8 (2013), 1426–1438.

[88] Görlitz, L., Menze, B. H., Weber, M.-A., Kelm, B. M., and Hamprecht, F. A. Semi-supervised tumor detection in magnetic resonance spectroscopic images using discriminative random fields. In *Pattern Recognition*. Springer, 2007, pp. 224–233.

[89] Govindaraju, V., Young, K., and Maudsley, A. A. Proton NMR chemical shifts and coupling constants for brain metabolites. *NMR in Biomedicine 13*, 3 (2000), 129–153.

[90] Grant, R. Overview: brain tumour diagnosis and management/Royal College of Physicians guidelines. *Journal of Neurology, Neurosurgery & Psychiatry 75*, suppl 2 (2004), ii18–ii23.

[91] Graves, E. E., Nelson, S. J., Vigneron, D. B., Verhey, L., McDermott, M., Larson, D., Chang, S., Prados, M. D., and Dillon, W. P. Serial proton MR spectroscopic imaging of recurrent malignant gliomas after gamma knife radiosurgery. *American Journal of Neuroradiology 22*, 4 (2001), 613–624.

[92] Guo, A. C., Cummings, T. J., Dash, R. C., and Provenzale, J. M. Lymphomas and high-grade astrocytomas: Comparison of water diffusibility and histologic characteristics 1. *Radiology 224*, 1 (2002), 177–183.

[93] Haase, A., Frahm, J., Hanicke, W., and Matthaei, D. 1H NMR chemical shift selective (CHESS) imaging. *Physics in Medicine and Biology 30*, 4 (1985), 341.

[94] Haeck, T., Maes, F., and Suetens, P. Automated model-based segmentation of brain tumors in MR images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2015), Springer, pp. 25–28.

[95] Hall, W. A., Martin, A., Liu, H., and Truwit, C. L. Improving diagnostic yield in brain biopsy: Coupling spectroscopic targeting with

real-time needle placement. *Journal of Magnetic Resonance imaging 13*, 1 (2001), 12–15.

[96] Hamamci, A., Kucuk, N., Karaman, K., Engin, K., and Unal, G. Tumor-cut: segmentation of brain tumors on contrast enhanced MR images for radiosurgery applications. *IEEE Transactions on Medical Imaging 31*, 3 (2012), 790–804.

[97] Hammoud, M. A., Sawaya, R., Shi, W., Thall, P. F., and Leeds, N. E. Prognostic significance of preoperative MRI scans in glioblastoma multiforme. *Journal of Neuro-Oncology 27*, 1 (1996), 65–73.

[98] Hattingen, E., Delic, O., Franz, K., Pilatus, U., Raab, P., Lanfermann, H., and Gerlach, R. 1H MRSI and progression-free survival in patients with WHO grades ii and iii gliomas. *Neurological Research 32*, 6 (2010), 593–602.

[99] Havaei, M., Jodoin, P.-M., and Larochelle, H. Efficient interactive brain tumor segmentation as within-brain kNN classification. In *ICPR* (2014), pp. 556–561.

[100] Hein, P. A., Eskey, C. J., Dunn, J. F., and Hug, E. B. Diffusion-weighted imaging in the follow-up of treated high-grade gliomas: tumor recurrence versus radiation injury. *American Journal of Neuroradiology 25*, 2 (2004), 201–209.

[101] Heinrich, K. E., Berry, M. W., and Homayouni, R. Gene tree labeling using nonnegative matrix factorization on biomedical literature. *Computational Intelligence and Neuroscience 2008* (2008), 2.

[102] Ho, S., Bullitt, E., and Gerig, G. Level-set evolution with region competition: automatic 3-D segmentation of brain tumors. In *International Conference on Pattern Recognition* (2002), vol. 1, IEEE, pp. 532–535.

[103] Hourani, R., Brant, L., Rizk, T., Weingart, J., Barker, P., and Horska, A. Can proton MR spectroscopic and perfusion imaging differentiate between neoplastic and nonneoplastic brain lesions in adults? *American Journal of Neuroradiology 29*, 2 (2008), 366–372.

[104] Hsieh, T. M., Liu, Y.-M., Liao, C.-C., Xiao, F., Chiang, I.-J., and Wong, J.-M. Automatic segmentation of meningioma from non-contrasted brain MRI integrating fuzzy clustering and region growing. *BMC Medical Informatics and Decision Making 11*, 1 (2011), 1.

[105] Hui, E. S., Cheung, M. M., Qi, L., and Wu, E. X. Towards better MR characterization of neural tissues using directional diffusion kurtosis analysis. *Neuroimage 42*, 1 (2008), 122–134.

[106] Ibanez, L., Schroeder, W., Ng, L., and Cates, J. The ITK software guide.

[107] Ion-Margineanu, A., Van Cauter, S., Sima, D. M., Maes, F., Van Gool, S. W., Sunaert, S., Himmelreich, U., and Van Huffel, S. Tumour relapse prediction using multiparametric MR data recorded during follow-up of GBM patients. *BioMed Research International 2015* (2015).

[108] Irfanoglu, M. O., Walker, L., Sarlls, J., Marenco, S., and Pierpaoli, C. Effects of image distortions originating from susceptibility variations and concomitant fields on diffusion MRI tractography results. *Neuroimage 61*, 1 (2012), 275–288.

[109] Jacobs, A. H., Kracht, L. W., Gossmann, A., Rüger, M. A., Thomas, A. V., Thiel, A., and Herholz, K. Imaging in neurooncology. *NeuroRx 2*, 2 (2005), 333–347.

[110] Jain, S., Sima, D., Sanaei Nezhad, F., Williams, S., Van Huffel, S., Maes, F., and Smeets, D. Patch based super-resolution of MR spectroscopic images. In *International Symposium on Biomedical Imaging* (2016), IEEE, pp. 452–456.

[111] Jellison, B. J., Field, A. S., Medow, J., Lazar, M., Salamat, M. S., and Alexander, A. L. Diffusion tensor imaging of cerebral white matter: a pictorial review of physics, fiber tract anatomy, and tumor imaging patterns. *American Journal of Neuroradiology 25*, 3 (2004), 356–369.

[112] Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., and Smith, S. M. FSL. *Neuroimage 62*, 2 (2012), 782–790.

[113] Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H., and Kaczynski, K. Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine 53*, 6 (2005), 1432–1440.

[114] Jensen, T. R., and Schmainda, K. M. Computer-aided detection of brain tumor invasion using multiparametric MRI. *Journal of Magnetic Resonance Imaging 30*, 3 (2009), 481–489.

[115] Juan-Albarracín, J., Fuster-Garcia, E., Manjón, J. V., Robles, M., Aparici, F., Martí-Bonmatí, L., and García-Gómez, J. M. Automated glioblastoma segmentation based on a multiparametric structured unsupervised classification. *PloS One 10*, 5 (2015), e0125143.

[116] Kallenberg, K., Goldmann, T., Menke, J., Strik, H., Bock, H., Stockhammer, F., Buhk, J., Frahm, J., Dechent, P., and Knauth, M. Glioma infiltration of the corpus callosum: early signs detected by DTI. *Journal of Neuro-Oncology 112*, 2 (2013), 217–222.

[117] Kassner, A., Annesley, D., Zhu, X., Li, K., Kamaly-Asl, I., Watson, Y., and Jackson, A. Abnormalities of the contrast re-circulation phase in cerebral tumors demonstrated using dynamic susceptibility contrast-enhanced imaging: A possible marker of vascular tortuosity. *Journal of Magnetic Resonance Imaging 11*, 2 (2000), 103–113.

[118] Kassner, A., and Thornhill, R. Texture analysis: a review of neurologic MR imaging applications. *American Journal of Neuroradiology 31*, 5 (2010), 809–816.

[119] Kazerooni, A. F., Mohseni, M., Rezaei, S., Bakhshandehpour, G., and Rad, H. S. Multi-parametric (ADC/PWI/T2-w) image fusion approach for accurate semi-automatic segmentation of tumorous regions in glioblastoma multiforme. *Magnetic Resonance Materials in Physics, Biology and Medicine 28*, 1 (2015), 13–22.

[120] Kim, H., and Park, H. Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method. *SIAM Journal on Matrix Analysis and Applications 30*, 2 (2008), 713–730.

[121] Kitis, O., Altay, H., Calli, C., Yunten, N., Akalin, T., and Yurtseven, T. Minimum apparent diffusion coefficients in the evaluation of brain tumors. *European Journal of Radiology 55*, 3 (2005), 393–400.

[122] Knopp, E. A., Cha, S., Johnson, G., Mazumdar, A., Golfinos, J. G., Zagzag, D., Miller, D. C., Kelly, P. J., and Kricheff, I. I. Glial neoplasms: Dynamic contrast-enhanced T2*-weighted MR imaging 1. *Radiology 211*, 3 (1999), 791–798.

[123] Koay, C. G., Özarslan, E., and Basser, P. J. A signal transformational framework for breaking the noise floor and its applications in MRI. *Journal of Magnetic Resonance 197*, 2 (2009), 108–119.

[124] Kohler, B. A., Ward, E., McCarthy, B. J., Schymura, M. J., Ries, L. A., Eheman, C., Jemal, A., Anderson, R. N., Ajani,

U. A., and Edwards, B. K. Annual report to the nation on the status of cancer, 1975–2007, featuring tumors of the brain and other nervous system. *Journal of the National Cancer Institute* (2011).

[125] Kono, K., Inoue, Y., Nakayama, K., Shakudo, M., Morino, M., Ohata, K., Wakasa, K., and Yamada, R. The role of diffusion-weighted imaging in patients with brain tumors. *American Journal of Neuroradiology 22*, 6 (2001), 1081–1088.

[126] Kounelakis, M. G., Dimou, I. N., Zervakis, M. E., Tsougos, I., Tsolaki, E., Kousi, E., Kapsalaki, E., and Theodorou, K. Strengths and weaknesses of 1.5T and 3T MRS data in brain glioma classification. *IEEE Transactions on Information Technology in Biomedicine 15*, 4 (2011), 647–654.

[127] Krawczyk, B. Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence* (2016), 1–12.

[128] Kuang, D., and Park, H. Fast rank-2 nonnegative matrix factorization for hierarchical document clustering. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2013), ACM, pp. 739–747.

[129] Kwon, D., Shinohara, R. T., Akbari, H., and Davatzikos, C. Combining generative models for multifocal glioma segmentation and registration. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2014), Springer, pp. 763–770.

[130] Lacroix, M., Abi-Said, D., Fourney, D. R., Gokaslan, Z. L., Shi, W., DeMonte, F., Lang, F. F., McCutcheon, I. E., Hassenbusch, S. J., Holland, E., Hess, K., Michael, C., Miller, D., and Sawaya, R. A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival. *Journal of Neurosurgery 95*, 2 (2001), 190–198.

[131] Lafferty, J., McCallum, A., and Pereira, F. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *International Conference on Machine Learning, ICML* (2001), vol. 1, pp. 282–289.

[132] Langville, A. N., Meyer, C. D., Albright, R., Cox, J., and Duling, D. Initializations for the nonnegative matrix factorization. In *Proceedings of the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2006), Citeseer, pp. 23–26.

[133] Laniado, M., Weinmann, H., Schörner, W., Felix, R., and Speck, U. First use of GdDTPA/dimeglumine in man. *Physiological chemistry and physics and medical NMR 16*, 2 (1983), 157–165.

[134] Laudadio, T., Croitor Sava, A., Sima, D., Wright, A., Heerschap, A., and Van Huffel, S. Hierarchical non-negative matrix factorization applied to in vivo 3T MRSI prostate data for automatic detection and visualization of tumours. In *Proceedings of the European Society for Magnetic Resonance in Medicine and Biology Annual meeting 2013* (2013), pp. 474–475.

[135] Laudadio, T., Martínez-Bisbal, M. C., Celda, B., and Van Huffel, S. Fast nosological imaging using canonical correlation analysis of brain data obtained by two-dimensional turbo spectroscopic imaging. *NMR in Biomedicine 21*, 4 (2008), 311–321.

[136] Laudadio, T., Mastronardi, N., Vanhamme, L., Van Hecke, P., and Van Huffel, S. Improved Lanczos algorithms for blackbox MRS data quantitation. *Journal of Magnetic Resonance 157*, 2 (2002), 292–297.

[137] Law, M., Cha, S., Knopp, E. A., Johnson, G., Arnett, J., and Litt, A. W. High-grade gliomas and solitary metastases: Differentiation by using perfusion and proton spectroscopic MR imaging 1. *Radiology 222*, 3 (2002), 715–721.

[138] Law, M., Oh, S., Babb, J. S., Wang, E., Inglese, M., Zagzag, D., Knopp, E. A., and Johnson, G. Low-grade gliomas: Dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging—prediction of patient clinical response 1. *Radiology 238*, 2 (2006), 658–667.

[139] Law, M., Yang, S., Babb, J. S., Knopp, E. A., Golfinos, J. G., Zagzag, D., and Johnson, G. Comparison of cerebral blood volume and vascular permeability from dynamic susceptibility contrast-enhanced perfusion MR imaging with glioma grade. *American Journal of Neuroradiology 25*, 5 (2004), 746–755.

[140] Law, M., Yang, S., Wang, H., Babb, J. S., Johnson, G., Cha, S., Knopp, E. A., and Zagzag, D. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. *American Journal of Neuroradiology 24*, 10 (2003), 1989–1998.

[141] Lawson, C. L., and Hanson, R. J. *Solving least squares problems*, vol. 161. SIAM, 1974.

[142] Lazar, M., Weinstein, D. M., Tsuruda, J. S., Hasan, K. M., Arfanakis, K., Meyerand, M. E., Badie, B., Rowley, H. A., Haughton, V., Field, A., et al. White matter tractography using diffusion tensor deflection. *Human Brain Mapping 18*, 4 (2003), 306–321.

[143] Le Bihan, D. Apparent diffusion coefficient and beyond: what diffusion MR imaging can tell us about tissue structure. *Radiology 268*, 2 (2013), 318–322.

[144] Le Bihan, D., and Basser, P. J. Molecular diffusion and nuclear magnetic resonance. *Diffusion and perfusion magnetic resonance imaging* (1995), 5–17.

[145] Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., and Laval-Jeantet, M. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology 161*, 2 (1986), 401–407.

[146] Le Bihan, D., Mangin, J.-F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., and Chabriat, H. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging 13*, 4 (2001), 534–546.

[147] Lee, C.-H., Wang, S., Murtha, A., Brown, M. R., and Greiner, R. Segmenting brain tumors using pseudo–conditional random fields. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2008), Springer, pp. 359–366.

[148] Lee, D. D., and Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature 401*, 6755 (1999), 788–791.

[149] Leemans, A., Jeurissen, B., Sijbers, J., and Jones, D. ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data. In *17th Annual Meeting of Intl Soc Mag Reson Med* (2009), vol. 209, p. 3537.

[150] Lev, M. H., Ozsunar, Y., Henson, J. W., Rasheed, A. A., Barest, G. D., Harsh, G. R., Fitzek, M. M., Chiocca, E. A., Rabinov, J. D., Csavoy, A. N., et al. Glial tumor grading and outcome prediction using dynamic spin-echo MR susceptibility mapping compared with conventional contrast-enhanced MR: confounding effect of elevated rCBV of oligodendroglimoas. *American Journal of Neuroradiology 25*, 2 (2004), 214–221.

[151] Levitt, M. H. *Spin dynamics: basics of nuclear magnetic resonance*. John Wiley & Sons, 2001.

[152] Li, S. Z. *Markov random field modeling in image analysis.* Springer Science & Business Media, 2009.

[153] Li, Y., Sima, D. M., Cauter, S. V., Croitor Sava, A. R., Himmelreich, U., Pi, Y., and Van Huffel, S. Hierarchical non-negative matrix factorization (hNMF): a tissue pattern differentiation method for glioblastoma multiforme diagnosis using MRSI. *NMR in Biomedicine 26*, 3 (2013), 307–319.

[154] Li, Y., Sima, D. M., Van Cauter, S., Himmelreich, U., Sava, A. R. C., Pi, Y., Liu, Y., and Van Huffel, S. Unsupervised nosologic imaging for glioma diagnosis. *IEEE Transactions on Biomedical Engineering 60*, 6 (2013), 1760–1763.

[155] Londoño, A., Castillo, M., Armao, D., Kwock, L., and Suzuki, K. Unusual MR spectroscopic imaging pattern of an astrocytoma: lack of elevated choline and high myo-inositol and glycine levels. *American Journal of Neuroradiology 24*, 5 (2003), 942–945.

[156] Louis, D. N., Holland, E. C., and Cairncross, J. G. Glioma classification: a molecular reappraisal. *The American Journal of Pathology 159*, 3 (2001), 779–786.

[157] Louis, D. N., Ohgaki, H., Wiestler, O. D., Cavenee, W. K., Burger, P. C., Jouvet, A., Scheithauer, B. W., and Kleihues, P. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathologica 114*, 2 (2007), 97–109.

[158] Lupo, J. M., Cha, S., Chang, S. M., and Nelson, S. J. Dynamic susceptibility-weighted perfusion imaging of high-grade gliomas: characterization of spatial heterogeneity. *American Journal of Neuroradiology 26*, 6 (2005), 1446–1454.

[159] Luts, J., Laudadio, T., Idema, A. J., Simonetti, A. W., Heerschap, A., Vandermeulen, D., Suykens, J. A., and Van Huffel, S. Nosologic imaging of the brain: segmentation and classification using MRI and MRSI. *NMR in Biomedicine 22*, 4 (2009), 374–390.

[160] MacKay, D. J. *Information theory, inference and learning algorithms.* Cambridge University Press, 2003.

[161] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging 16*, 2 (1997), 187–198.

[162] Maia, A. C., Malheiros, S. M., da Rocha, A. J., da Silva, C. J., Gabbai, A. A., Ferraz, F. A., and Stávale, J. N. MR cerebral blood volume maps correlated with vascular endothelial growth factor expression and tumor grade in nonenhancing gliomas. *American Journal of Neuroradiology 26*, 4 (2005), 777–783.

[163] Manoj, M., and Padmasuresh, L. An automated multimodal spectral cluster based segmentation for tumor and lesion detection in PET images. *Research Journal of Applied Sciences, Engineering and Technology 12*, 5 (2016), 522–527.

[164] Maudsley, A. A., Domenig, C., Govind, V., Darkazanli, A., Studholme, C., Arheart, K., and Bloomer, C. Mapping of brain metabolite distributions by volumetric proton MR spectroscopic imaging (MRSI). *Magnetic Resonance in Medicine 61*, 3 (2009), 548–559.

[165] Mazzara, G. P., Velthuizen, R. P., Pearlman, J. L., Greenberg, H. M., and Wagner, H. Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. *International Journal of Radiation Oncology* Biology* Physics 59*, 1 (2004), 300–312.

[166] McInerney, T., and Terzopoulos, D. Deformable models in medical image analysis: a survey. *Medical Image Analysis 1*, 2 (1996), 91–108.

[167] McKnight, T. R., Lamborn, K. R., Love, T. D., Berger, M. S., Chang, S., Dillon, W. P., Bollen, A., and Nelson, S. J. Correlation of magnetic resonance spectroscopic and growth characteristics within grades ii and iii gliomas. *Journal of Neurosurgery 106*, 4 (2007), 660–666.

[168] Menze, B., Reyes, M., Van Leemput, K., et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging 34*, 10 (2015), 1993–2024.

[169] Menze, B. H., Van Leemput, K., Lashkari, D., Weber, M.-A., Ayache, N., and Golland, P. A generative model for brain tumor segmentation in multi-modal images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 13. 2010, pp. 151–159.

[170] Mescher, M., Merkle, H., Kirsch, J., Garwood, M., and Gruetter, R. Simultaneous in vivo spectral editing and water suppression. *NMR in Biomedicine 11*, EPFL-ARTICLE-177509 (1998), 266–272.

[171] Mizutani, T. Robustness analysis of preconditioned successive projection algorithm for general form of separable NMF problem. *arXiv preprint arXiv:1506.08387* (2015).

[172] Moots, P. L., Maciunas, R. J., Eisert, D. R., Parker, R. A., Laporte, K., and Abou-Khalil, B. The course of seizure disorders in patients with malignant gliomas. *Archives of Neurology 52*, 7 (1995), 717–724.

[173] Mukherjee, P., Berman, J., Chung, S., Hess, C., and Henry, R. Diffusion tensor MR imaging and fiber tractography: theoretic underpinnings. *American Journal of Neuroradiology 29*, 4 (2008), 632–641.

[174] Muti, M., Aprile, I., Principi, M., Italiani, M., Guiducci, A., Giulianelli, G., and Ottaviano, P. Study on the variations of the apparent diffusion coefficient in areas of solid tumor in high grade gliomas. *Magnetic Resonance Imaging 20*, 9 (2002), 635–641.

[175] Nie, J., Xue, Z., Liu, T., Young, G. S., Setayesh, K., Guo, L., and Wong, S. T. Automated brain tumor segmentation using spatial accuracy-weighted hidden Markov random field. *Computerized Medical Imaging and Graphics 33*, 6 (2009), 431–441.

[176] Niegowski, M., and Zivanovic, M. ECG-EMG separation by using enhanced non-negative matrix factorization. In *36th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society (EMBC)* (2014), IEEE, pp. 4212–4215.

[177] Nocedal, J., and Wright, S. J. *Numerical Optimization, 2nd Edition*. Springer, 2006.

[178] Noch, E., and Khalili, K. Molecular mechanisms of necrosis in glioblastoma: the role of glutamate excitotoxicity. *Cancer Biology & Therapy 8*, 19 (2009), 1791–1797.

[179] Nock, R., and Nielsen, F. On weighting clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence 28*, 8 (2006), 1223–1235.

[180] Ogg, R. J., Kingsley, R., and Taylor, J. S. WET, a T1-and B1-insensitive water-suppression method for in vivo localized 1H NMR spectroscopy. *Journal of Magnetic Resonance, Series B 104*, 1 (1994), 1–10.

[181] Ortega-Martorell, S., Lisboa, P. J., Vellido, A., Julià-Sapé, M., and Arús, C. Non-negative matrix factorisation methods for the spectral decomposition of MRS data from human brain tumours. *BMC Bioinformatics 13*, 1 (2012), 38.

[182] Ortega-Martorell, S., Lisboa, P. J., Vellido, A., Simões, R. V., Pumarola, M., Julià-Sapé, M., and Arús, C. Convex non-negative matrix factorization for brain tumor delimitation from MRSI data. *PloS One 7*, 10 (2012), e47824.

[183] Østergaard, L., Weisskoff, R. M., Chesler, D. A., Gyldensted, C., and Rosen, B. R. High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. part i: Mathematical approach and statistical analysis. *Magnetic Resonance in Medicine 36*, 5 (1996), 715–725.

[184] Ostrom, Q. T., Gittleman, H., Liao, P., Rouse, C., Chen, Y., Dowling, J., Wolinsky, Y., Kruchko, C., and Barnholtz-Sloan, J. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2007–2011. *Neuro-Oncology 16*, suppl 4 (2014), iv1–iv63.

[185] Otazo, R., Mueller, B., Ugurbil, K., Wald, L., and Posse, S. Signal-to-noise ratio and spectral linewidth improvements between 1.5 and 7 Tesla in proton echo-planar spectroscopic imaging. *Magnetic Resonance in Medicine 56*, 6 (2006), 1200–1210.

[186] Paatero, P., and Tapper, U. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics 5*, 2 (1994), 111–126.

[187] Padhani, A. R., and Miles, K. A. Multiparametric imaging of tumor response to therapy 1. *Radiology 256*, 2 (2010), 348–364.

[188] Padole, V. B., and Chaudhari, D. Detection of brain tumor in MRI images using mean shift algorithm and normalized cut method. *International Journal of Engineering and Advanced Technology* (2012).

[189] Pauleit, D., Langen, K.-J., Floeth, F., Hautzel, H., Riemenschneider, M. J., Reifenberger, G., Shah, N. J., and Müller, H.-W. Can the apparent diffusion coefficient be used as a noninvasive parameter to distinguish tumor tissue from peritumoral tissue in cerebral gliomas? *Journal of Magnetic Resonance Imaging 20*, 5 (2004), 758–764.

[190] Paulus, W., and Peiffer, J. Intratumoral histologic heterogeneity of gliomas. A quantitative study. *Cancer 64*, 2 (1989), 442–447.

[191] Perthen, J. E., Calamante, F., Gadian, D. G., and Connelly, A. Is quantification of bolus tracking MRI reliable without deconvolution? *Magnetic Resonance in Medicine 47*, 1 (2002), 61–67.

[192] Phillips, W., Velthuizen, R., Phuphanich, S., Hall, L., Clarke, L., and Silbiger, M. Application of fuzzy c-means segmentation technique for tissue differentiation in MR images of a hemorrhagic glioblastoma multiforme. *Magnetic Resonance Imaging 13*, 2 (1995), 277–290.

[193] Pirzkall, A., McKnight, T. R., Graves, E. E., Carol, M. P., Sneed, P. K., Wara, W. W., Nelson, S. J., Verhey, L. J., and Larson, D. A. MR-spectroscopy guided target delineation for high-grade gliomas. *International Journal of Radiation Oncology* Biology* Physics 50*, 4 (2001), 915–928.

[194] Plotkin, M., Eisenacher, J., Bruhn, H., Wurm, R., Michel, R., Stockhammer, F., Feussner, A., Dudeck, O., Wust, P., Felix, R., et al. 123I-IMT SPECT and 1HMR-spectroscopy at 3.0T in the differential diagnosis of recurrent or residual gliomas: a comparative study. *Journal of Neuro-Oncology 70*, 1 (2004), 49–58.

[195] Podo, F. Tumour phospholipid metabolism. *NMR in Biomedicine 12*, 7 (1999), 413–439.

[196] Poot, D. H., den Dekker, A. J., Achten, E., Verhoye, M., and Sijbers, J. Optimal experimental design for diffusion kurtosis imaging. *IEEE Transactions on Medical Imaging 29*, 3 (2010), 819–829.

[197] Posner, J. Management of brain metastases. *Revue Neurologique 148*, 6-7 (1991), 477–487.

[198] Poullet, J.-B. *Quantification and classification of magnetic resonance spectroscopic data for brain tumor diagnosis*. PhD thesis, Leuven, Belgium, 2008.

[199] Poullet, J.-B., Sima, D. M., Simonetti, A. W., De Neuter, B., Vanhamme, L., Lemmerling, P., and Van Huffel, S. An automated quantitation of short echo time MRS spectra in an open source software environment: AQSES. *NMR in Biomedicine 20*, 5 (2007), 493–504.

[200] Poullet, J.-B., Sima, D. M., Van Huffel, S., and Van Hecke, P. Frequency-selective quantitation of short-echo time 1H magnetic resonance spectra. *Journal of Magnetic Resonance 186*, 2 (2007), 293–304.

[201] Prastawa, M., Bullitt, E., and Gerig, G. Simulation of brain tumors in MR images for evaluation of segmentation efficacy. *Medical Image Analysis 13*, 2 (2009), 297–311.

[202] Prastawa, M., Bullitt, E., Moon, N., Van Leemput, K., and Gerig, G. Automatic brain tumor segmentation by subject specific modification of atlas priors 1. *Academic Radiology 10*, 12 (2003), 1341–1348.

[203] Price, S., Jena, R., Burnet, N., Hutchinson, P., Dean, A., Pena, A., Pickard, J., Carpenter, T., and Gillard, J. Improved delineation of glioma margins and regions of infiltration with the use of diffusion tensor imaging: an image-guided biopsy study. *American Journal of Neuroradiology 27*, 9 (2006), 1969–1974.

[204] Provencher, S. W. Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magnetic Resonance in Medicine 30*, 6 (1993), 672–679.

[205] Qiao, H. New SVD based initialization strategy for non-negative matrix factorization. *Pattern Recognition Letters 63* (2014), 71–77.

[206] Raab, P., Hattingen, E., Franz, K., Zanella, F. E., and Lanfermann, H. Cerebral gliomas: Diffusional kurtosis imaging analysis of microstructural differences 1. *Radiology 254*, 3 (2010), 876–881.

[207] Ratiney, H., Coenradie, Y., Cavassila, S., van Ormondt, D., and Graveron-Demilly, D. Time-domain quantitation of 1H short echo-time signals: background accommodation. *Magnetic Resonance Materials in Physics, Biology and Medicine 16*, 6 (2004), 284–296.

[208] Ratiney, H., Sdika, M., Coenradie, Y., Cavassila, S., Ormondt, D. v., and Graveron-Demilly, D. Time-domain semi-parametric estimation based on a metabolite basis set. *NMR in Biomedicine 18*, 1 (2005), 1–13.

[209] Reardon, D. A., Rich, J. N., Friedman, H. S., and Bigner, D. D. Recent advances in the treatment of malignant astrocytoma. *Journal of Clinical Oncology 24*, 8 (2006), 1253–1265.

[210] Rexilius, J., Hahn, H. K., Klein, J., Lentschig, M. G., and Peitgen, H.-O. Multispectral brain tumor segmentation based on histogram model adaptation. In *Medical Imaging* (2007), International Society for Optics and Photonics, p. 65140V.

[211] Rezaei, M., Boostani, R., and Rezaei, M. An efficient initialization method for nonnegative matrix factorization. *Journal of Applied Sciences 11*, 2 (2011), 354–359.

[212] Ricci, R., Bacci, A., Tugnoli, V., Battaglia, S., Maffei, M., Agati, R., and Leonardi, M. Metabolic findings on 3T 1H-MR spectroscopy in peritumoral brain edema. *American Journal of Neuroradiology 28*, 7 (2007), 1287–1291.

[213] Rorden, C., Karnath, H.-O., and Bonilha, L. Improving lesion-symptom mapping. *Journal of Cognitive Neuroscience 19*, 7 (2007), 1081–1088.

[214] Rosen, B. R., Belliveau, J. W., Vevea, J. M., and Brady, T. J. Perfusion imaging with NMR contrast agents. *Magnetic Resonance in Medicine 14*, 2 (1990), 249–265.

[215] Ross, A. J., and Sachdev, P. S. Magnetic resonance spectroscopy in cognitive research. *Brain Research Reviews 44*, 2 (2004), 83–102.

[216] Ruan, S., Lebonvallet, S., Merabet, A., and Constans, J.-M. Tumor segmentation from a multispectral MRI images by using support vector machine classification. In *Proceedings of the Fourth IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2007), IEEE, pp. 1236–1239.

[217] Sadeghi, N., D'Haene, N., Decaestecker, C., Levivier, M., Metens, T., Maris, C., Wikler, D., Balériaux, D., Salmon, I., and Goldman, S. Apparent diffusion coefficient and cerebral blood volume in brain gliomas: relation to tumor cell density and tumor microvessel density based on stereotactic biopsies. *American Journal of Neuroradiology 29*, 3 (2008), 476–482.

[218] Saha, B. N., Ray, N., Greiner, R., Murtha, A., and Zhang, H. Quick detection of brain tumors and edemas: A bounding box method using symmetry. *Computerized Medical Imaging and Graphics 36*, 2 (2012), 95–107.

[219] Sajda, P., Du, S., Brown, T. R., Stoyanova, R., Shungu, D. C., Mao, X., and Parra, L. C. Nonnegative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain. *IEEE Transactions on Medical Imaging 23*, 12 (2004), 1453–1465.

[220] Sauwen, N., Acou, M., Bharath, H., Sima, D., Veraart, J., Maes, F., Himmelreich, U., Achten, E., and Van Huffel, S. The successive projection algorithm as an initialization method for brain tumor segmentation using non-negative matrix factorization. Submitted as a journal paper for peer review.

[221] Sauwen, N., Acou, M., Bharath, H., Sima, D., Veraart, J., Maes, F., Himmelreich, U., Achten, E., and Van Huffel, S. Initializing nonnegative matrix factorization using the successive projection algorithm for multi-parametric medical image segmentation. In *Proceedings of the 24th ESANN European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning* (2016), pp. 265–270.

[222] Sauwen, N., Acou, M., Sima, D., Maes, F., Himmelreich, U., Achten, E., and Van Huffel, S. A semi-automated segmentation framework for MRI based brain tumor segmentation using regularized nonnegative matrix factorization. In *International Conference on Signal Image Technology and Internet-based Systems* (2016), Springer. Accepted for publication.

[223] Sauwen, N., Acou, M., Sima, D., Veraart, J., Maes, F., Himmelreich, U., Achten, E., and Van Huffel, S. Semi-automated brain tumor segmentation on multi-parametric MRI using regularized non-negative matrix factorization. Submitted as a journal paper for peer review.

[224] Sauwen, N., Acou, M., Van Cauter, S., Sima, D., Veraart, J., Maes, F., Himmelreich, U., Achten, E., and Van Huffel, S. Comparison of unsupervised classification methods for brain tumor segmentation using multi-parametric MRI. *NeuroImage: Clinical 12* (2016), 753–764.

[225] Sauwen, N., Sima, D. M., Van Cauter, S., Veraart, J., Leemans, A., Maes, F., Himmelreich, U., and Van Huffel, S. Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using multi-parametric MRI. *NMR in Biomedicine 28*, 12 (2015), 1599–1624.

[226] Schneider, C. A., Rasband, W. S., Eliceiri, K. W., et al. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods 9*, 7 (2012), 671–675.

[227] Schölkopf, B., and Smola, A. J. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2002.

[228] Schroff, F., Criminisi, A., and Zisserman, A. Object class segmentation using random forests. In *BMVC* (2008), pp. 1–10.

[229] Scott, J., Brasher, P., Sevick, R., Rewcastle, N., and Forsyth, P. How often are nonenhancing supratentorial gliomas malignant? a population study. *Neurology 59*, 6 (2002), 947–949.

[230] Server, A., Kulle, B., Gadmar, Ø. B., Josefsen, R., Kumar, T., and Nakstad, P. H. Measurements of diagnostic examination performance using quantitative apparent diffusion coefficient and proton MR spectroscopic imaging in the preoperative evaluation of tumor grade in cerebral gliomas. *European Journal of Radiology 80*, 2 (2011), 462–470.

[231] Shahnaz, F., Berry, M. W., Pauca, V. P., and Plemmons, R. J. Document clustering using nonnegative matrix factorization. *Information Processing & Management 42*, 2 (2006), 373–386.

[232] Shi, J., and Malik, J. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence 22*, 8 (2000), 888–905.

[233] Shiroishi, M. S., Castellazzi, G., Boxerman, J. L., D'Amore, F., Essig, M., Nguyen, T. B., Provenzale, J. M., Enterline, D. S., Anzalone, N., Dörfler, A., et al. Principles of T2*-weighted dynamic susceptibility contrast MRI technique in brain tumor imaging. *Journal of Magnetic Resonance Imaging 41*, 2 (2015), 296–313.

[234] Sinha, S., Bastin, M. E., Whittle, I. R., and Wardlaw, J. M. Diffusion tensor MR imaging of high-grade cerebral gliomas. *American Journal of Neuroradiology 23*, 4 (2002), 520–527.

[235] Sizoo, E. M., Braam, L., Postma, T. J., Pasman, H. R. W., Heimans, J. J., Klein, M., Reijneveld, J. C., and Taphoorn, M. J. Symptoms and problems in the end-of-life phase of high-grade glioma patients. *Neuro-Oncology 12*, 11 (2010), 1162–1166.

[236] Sladky, R., Friston, K. J., Tröstl, J., Cunnington, R., Moser, E., and Windischberger, C. Slice-timing effects and their correction in functional MRI. *Neuroimage 58*, 2 (2011), 588–594.

[237] Sled, J. G., Zijdenbos, A. P., and Evans, A. C. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging 17*, 1 (1998), 87–97.

[238] Smith, E. A., Carlos, R. C., Junck, L. R., Tsien, C. I., Elias, A., and Sundgren, P. C. Developing a clinical decision model: MR spectroscopy to differentiate between recurrent tumor and radiation change in patients with new contrast-enhancing lesions. *American Journal of Roentgenology 192*, 2 (2009), W45–W52.

[239] Smith, J. S., Cha, S., Mayo, M. C., McDermott, M. W., Parsa, A. T., Chang, S. M., Dillon, W. P., and Berger, M. S. Serial diffusion-weighted magnetic resonance imaging in cases of glioma:

distinguishing tumor recurrence from postresection injury. *Journal of Neurosurgery 103*, 3 (2005), 428–438.

[240] Srinivasan, R., Sailasuta, N., Hurd, R., Nelson, S., and Pelletier, D. Evidence of elevated glutamate in multiple sclerosis using magnetic resonance spectroscopy at 3T. *Brain 128*, 5 (2005), 1016–1025.

[241] Stadlbauer, A., Gruber, S., Nimsky, C., Fahlbusch, R., Hammen, T., Buslei, R., Tomandl, B., Moser, E., and Ganslandt, O. Preoperative grading of gliomas by using metabolite quantification with high-spatial-resolution proton MR spectroscopic imaging 1. *Radiology 238*, 3 (2006), 958–969.

[242] Stejskal, E. O., and Tanner, J. E. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics 42*, 1 (1965), 288–292.

[243] Stewart, G. N. Researches on the circulation time and on the influences which affect it. *The Journal of Physiology 22*, 3 (1897), 159.

[244] Stupp, R., Mason, W. P., Van Den Bent, M. J., Weller, M., Fisher, B., Taphoorn, M. J., Belanger, K., Brandes, A. A., Marosi, C., Bogdahn, U., et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New England Journal of Medicine 352*, 10 (2005), 987–996.

[245] Subbanna, N. K., Precup, D., Collins, D. L., and Arbel, T. Hierarchical probabilistic gabor and MRF segmentation of brain tumours in MRI volumes. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2013), Springer, pp. 751–758.

[246] Sugahara, T., Korogi, Y., Kochi, M., Ikushima, I., Hirai, T., Okuda, T., Shigematsu, Y., Liang, L., Ge, Y., Ushio, Y., et al. Correlation of MR imaging-determined cerebral blood volume maps with histologic and angiographic determination of vascularity of gliomas. *AJR. American Journal of Roentgenology 171*, 6 (1998), 1479–1486.

[247] Tabesh, A., Jensen, J. H., Ardekani, B. A., and Helpern, J. A. Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging. *Magnetic Resonance in Medicine 65*, 3 (2011), 823–836.

[248] Tofts, P. S., Brix, G., Buckley, D. L., Evelhoch, J. L., Henderson, E., Knopp, M. V., Larsson, H. B., Lee, T.-Y., Mayr, N. A., Parker, G. J., et al. Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer:

standardized quantities and symbols. *Journal of Magnetic Resonance Imaging 10*, 3 (1999), 223–232.

[249] Tong, Z., Yamaki, T., Harada, K., and Houkin, K. In vivo quantification of the metabolites in normal brain and brain tumors by proton MR spectroscopy using water as an internal standard. *Magnetic Resonance Imaging 22*, 7 (2004), 1017–1024.

[250] Tuceryan, M., Jain, A. K., et al. Texture analysis. *Handbook of Pattern Recognition and Computer Vision 2* (1993), 207–248.

[251] Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., and Gee, J. C. N4ITK: improved N3 bias correction. *IEEE Transactions on Medical Imaging 29*, 6 (2010), 1310–1320.

[252] Van Cauter, S., De Keyzer, F., Sima, D. M., Sava, A. C., D'Arco, F., Veraart, J., Peeters, R. R., Leemans, A., Van Gool, S., Wilms, G., et al. Integrating diffusion kurtosis imaging, dynamic susceptibility-weighted contrast-enhanced MRI, and short echo time chemical shift imaging for grading gliomas. *Neuro-Oncology 16*, 7 (2014), 1010–1021.

[253] Van Cauter, S., Veraart, J., Sijbers, J., Peeters, R. R., Himmelreich, U., De Keyzer, F., Van Gool, S. W., Van Calenbergh, F., De Vleeschouwer, S., Van Hecke, W., and Sunaert, S. Gliomas: diffusion kurtosis MR imaging in grading. *Radiology 263*, 2 (2012), 492–501.

[254] van den Bent, M. J. Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. *Acta Neuropathologica 120*, 3 (2010), 297–304.

[255] Van Gool, S., Maes, W., Ardon, H., Verschuere, T., Van Cauter, S., and De Vleeschouwer, S. Dendritic cell therapy of high-grade gliomas. *Brain Pathology 19*, 4 (2009), 694–712.

[256] Vavasis, S. A. On the complexity of nonnegative matrix factorization. *SIAM Journal on Optimization 20*, 3 (2009), 1364–1377.

[257] Veraart, J., Poot, D. H., Van Hecke, W., Blockx, I., Van der Linden, A., Verhoye, M., and Sijbers, J. More accurate estimation of diffusion tensor parameters using diffusion kurtosis imaging. *Magnetic Resonance in Medicine 65*, 1 (2011), 138–145.

[258] Veraart, J., Sijbers, J., Sunaert, S., Leemans, A., and Jeurissen, B. Weighted linear least squares estimation of diffusion MRI parameters: strengths, limitations, and pitfalls. *NeuroImage 81* (2013), 335–346.

[259] Verma, R., Zacharaki, E. I., Ou, Y., Cai, H., Chawla, S., Lee, S.-K., Melhem, E. R., Wolf, R., and Davatzikos, C. Multiparametric tissue characterization of brain neoplasms and their recurrence using pattern classification of MR images. *Academic Radiology 15*, 8 (2008), 966–977.

[260] Vervliet, N., Debals, O., Sorber, L., Van Barel, M., and De Lathauwer, L. Tensorlab v3.0 (March 2016). `http://www.tensorlab.net`. [Online; accessed 24-June-2016].

[261] Villringer, A., Rosen, B. R., Belliveau, J. W., Ackerman, J. L., Lauffer, R. B., Buxton, R. B., Chao, Y.-S., Wedeenand, V. J., and Brady, T. J. Dynamic imaging with lanthanide chelates in normal brain: contrast due to magnetic susceptibility effects. *Magnetic Resonance in Medicine 6*, 2 (1988), 164–174.

[262] Von Luxburg, U. A tutorial on spectral clustering. *Statistics and Computing 17*, 4 (2007), 395–416.

[263] Vovk, U., Pernus, F., and Likar, B. A review of methods for correction of intensity inhomogeneity in MRI. *IEEE Transactions on Medical Imaging 26*, 3 (2007), 405–421.

[264] Wagner, M., Nafe, R., Jurcoane, A., Pilatus, U., Franz, K., Rieger, J., Steinbach, J. P., and Hattingen, E. Heterogeneity in malignant gliomas: a magnetic resonance analysis of spatial distribution of metabolite changes and regional blood volume. *Journal of Neuro-Oncology 103*, 3 (2011), 663–672.

[265] Walker, M. D., Alexander Jr, E., Hunt, W. E., MacCarty, C. S., Mahaley Jr, M. S., Mealey Jr, J., Norrell, H. A., Owens, G., Ransohoff, J., Wilson, C. B., et al. Evaluation of BCNU and/or radiotherapy in the treatment of anaplastic gliomas: a cooperative clinical trial. *Journal of neurosurgery 49*, 3 (1978), 333–343.

[266] Wang, T., Cheng, I., Basu, A., et al. Fluid vector flow and applications in brain tumor segmentation. *IEEE Transactions on Biomedical Engineering 56*, 3 (2009), 781–789.

[267] Wang, Y.-X., and Zhang, Y.-J. Nonnegative matrix factorization: A comprehensive review. *IEEE Transactions on Knowledge and Data Engineering 25*, 6 (2013), 1336–1353.

[268] Weber, M.-A., Henze, M., Tüttenberg, J., Stieltjes, B., Meissner, M., Zimmer, F., Burkholder, I., Kroll, A., Combs, S. E., Vogt-Schaden, M., et al. Biopsy targeting gliomas: do functional imaging techniques identify similar target areas? *Investigative Radiology 45*, 12 (2010), 755–768.

[269] Weisskoff, R., Belliveau, J., Kwong, K., and Rosen, B. Functional MR imaging of capillary hemodynamics. *Magnetic Resonance Angiography: Concepts and Applications. St. Louis, Mo: Mosby-Year Book* (1993), 473–484.

[270] Weisskoff, R., Boxerman, J., Sorensen, A., Kulke, S., Campbell, T., and Rosen, B. Simultaneous blood volume and permeability mapping using a single Gd-based contrast injection. In *Proceedings of the Society of Magnetic Resonance, Second Annual Meeting* (1994), pp. 6–12.

[271] Weltens, C., Menten, J., Feron, M., Bellon, E., Demaerel, P., Maes, F., Van den Bogaert, W., and van der Schueren, E. Interobserver variations in gross tumor volume delineation of brain tumors on computed tomography and impact of magnetic resonance imaging. *Radiotherapy and Oncology 60*, 1 (2001), 49–59.

[272] Wen, P. Y., Macdonald, D. R., Reardon, D. A., Cloughesy, T. F., Sorensen, A. G., Galanis, E., DeGroot, J., Wick, W., Gilbert, M. R., Lassman, A. B., et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *Journal of Clinical Oncology 28*, 11 (2010), 1963–1972.

[273] Wesseling, P., van der Laak, J. A., Link, M., Teepen, H. L., and Ruiter, D. J. Quantitative analysis of microvascular changes in diffuse astrocytic neoplasms with increasing grade of malignancy. *Human Pathology 29*, 4 (1998), 352–358.

[274] Weybright, P., Sundgren, P. C., Maly, P., Hassan, D. G., Nan, B., Rohrer, S., and Junck, L. Differentiation between brain tumor recurrence and radiation injury using MR spectroscopy. *American Journal of Roentgenology 185*, 6 (2005), 1471–1476.

[275] Wild, S., Curry, J., and Dougherty, A. Improving non-negative matrix factorizations through structured initialization. *Pattern Recognition 37*, 11 (2004), 2217–2232.

[276] Willems, J., and Alva-Willems, J. Accuracy of cytologic diagnosis of central nervous system neoplasms in sterotactic biopsies. *Acta Cytologica 28*, 3 (1983), 243–249.

[277] Wilson, M., Reynolds, G., Kauppinen, R. A., Arvanitis, T. N., and Peet, A. C. A constrained least-squares approach to the automated quantitation of in vivo 1H magnetic resonance spectroscopy data. *Magnetic Resonance in Medicine 65*, 1 (2011), 1–12.

[278] Winston, G. P. The physical and biological basis of quantitative parameters derived from diffusion MRI. *Quantitative Imaging in Medicine and Surgery 2*, 4 (2012), 254–265.

[279] Wyss, M., and Kaddurah-Daouk, R. Creatine and creatinine metabolism. *Physiological Reviews 80*, 3 (2000), 1107–1213.

[280] Xie, Y., Ho, J., and Vemuri, B. C. Nonnegative factorization of diffusion tensor images and its applications. In *Information Processing in Medical Imaging* (2011), vol. 22, Springer, pp. 550–561.

[281] Yang, F., and Grigsby, P. Spectral clustering for FDG-PET cervical tumor segmentation. *International Journal of Radiation Oncology* Biology* Physics 78*, 3 (2010), S494.

[282] Yang, Z., Zhu, Z., and Oja, E. Automatic rank determination in projective nonnegative matrix factorization. In *International Conference on Latent Variable Analysis and Signal Separation* (2010), Springer, pp. 514–521.

[283] Young, G. S. Advanced MRI of adult brain tumors. *Neurologic Clinics 25*, 4 (2007), 947–973.

[284] Yu, C. S., Li, K. C., Xuan, Y., Ji, X. M., and Qin, W. Diffusion tensor tractography in patients with cerebral tumors: a helpful technique for neurosurgical planning and postoperative assessment. *European Journal of Radiology 56*, 2 (2005), 197–204.

[285] Yu, S., Zhang, Y., Liu, W., Zhao, N., Xiao, X., and Yin, G. A novel initialization method for nonnegative matrix factorization and its application in component recognition with three-dimensional fluorescence spectra. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 86* (2012), 315–319.

[286] Zacharaki, E. I., Shen, D., Lee, S.-K., and Davatzikos, C. ORBIT: a multiresolution framework for deformable registration of brain tumor images. *IEEE Transactions on Medical Imaging 27*, 8 (2008), 1003–1017.

[287] Zhang, J., Rivard, B., and Rogge, D. The successive projection algorithm (SPA), an algorithm with a spatial constraint for the automatic search of endmembers in hyperspectral data. *Sensors 8*, 2 (2008), 1321–1342.

[288] Zhang, Y., Brady, M., and Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging 20*, 1 (2001), 45–57.

[289] Zhao, L., Zhuang, G., and Xu, X. Facial expression recognition based on PCA and NMF. In *7th World Congress on Intelligent Control and Automation, 2008. WCICA 2008.* (2008), IEEE, pp. 6826–6829.

[290] Zheng, Z., Yang, J., and Zhu, Y. Initialization enhancer for non-negative matrix factorization. *Engineering Applications of Artificial Intelligence 20*, 1 (2007), 101–110.

[291] Zierler, K. L. Equations for measuring blood flow by external monitoring of radioisotopes. *Circulation Research 16*, 4 (1965), 309–321.

[292] Zikic, D., Glocker, B., Konukoglu, E., Criminisi, A., Demiralp, C., Shotton, J., Thomas, O., Das, T., Jena, R., and Price, S. Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR. In *International COnference on Medical Image Computing and Computer-Assisted Intervention*, vol. LCNS 7512. 2012, pp. 369–376.

[293] Zonari, P., Baraldi, P., and Crisi, G. Multimodal MRI in the characterization of glial neoplasms: the combined role of single-voxel MR spectroscopy, diffusion imaging and echo-planar perfusion imaging. *Neuroradiology 49*, 10 (2007), 795–803.

[294] Zou, K. H., Wells, W. M., Kikinis, R., and Warfield, S. K. Three validation metrics for automated probabilistic image segmentation of brain tumours. *Statistics in Medicine 23*, 8 (2004), 1259–1282.

Curriculum vitae

Nicolas Sauwen was born in Wilrijk, Belgium on February 3, 1983. He received the M.Sc. degree in mechanical engineering from KU Leuven (Belgium) in September 2006. During the last 2 years of his studies, he specialized in biomechanical engineering. After his studies, he was employed by the KU Leuven Biomechanical Engineering department as a scientific researcher. He studied the mechanical stability of uncemented hip prostheses in continuation of his master thesis entitled 'Prediction of hip implant fixation based on pre-operative CT imaging'.

From September 2007 till May 2011, he worked at the Customer Services Test division of LMS International, where he assisted customers in the usage of noise and vibration measurement equipment and software. From June 2011 till October 2011, he was involved in a micro-credits project as a volunteering worker in Tamale, Ghana by the Kidz Active foundation (The Netherlands).

In October 2012 he joined the KU Leuven Electrical Engineering department (ESAT-STADIUS) as a PhD student under the supervision of Prof. Sabine Van Huffel. His research has been part of a collaborative project involving KU Leuven, UZ Leuven and UZ Gent and was funded by the Research Foundation - Flanders (FWO project G.0869.12N).

List of publications

Papers in international journals

1. **Sauwen, N.**, Acou, M., Van Cauter, S., Sima, D.M., Veraart, J., Maes, F., Himmelreich, U., Achten, E. and Van Huffel, S. Comparison of unsupervised classification methods for brain tumor segmentation using multi-parametric MRI. NeuroImage: Clinical 12 (2016), 753-764.

2. Bharath, H.N., Sima, D.M., **Sauwen, N.**, Himmelreich, U., De Lathauwer, L., and Van Huffel, S. Non-Negative Canonical Polyadic Decomposition for Tissue Type Differentiation in Gliomas. IEEE Journal of Biomedical Health Informatics. Epub ahead of print.

3. **Sauwen, N.**, Sima, D.M., Van Cauter, S., Veraart, J., Leemans, A., Maes, F., Himmelreich, U., and Van Huffel, S. Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using multi-parametric MRI. NMR in Biomedicine 28, 12 (2015), 1599–1624.

Papers submitted to international journals

1. **Sauwen, N.**, Acou, M., Sima, D.M., Veraart, J., Maes, F., Himmelreich, U., Achten, E. and Van Huffel, S. Semi-automated brain tumor segmentation on multi-parametric MRI using regularized non-negative matrix factorization. BMC Medical Imaging, submitted 24 August, 2016.

2. **Sauwen, N.**, Acou, M., Bharath, H.N., Sima, D.M., Veraart, J., Maes, F., Himmelreich, U., Achten, E. and Van Huffel, S. The successive projection algorithm as an initialization method for brain tumor segmentation using non-negative matrix factorization. PlosOne, submitted 20 May, 2016.

Book chapters

1. Laudadio, T., Croitor Sava, A.R., Li, Y., **Sauwen, N.**, Sima, D.M., Van Huffel, S. 'NMF in MR spectroscopy', Chapter 7 in: Non-negative matrix factorization techniques. Advances in theory and applications. Springer, 2016, pp. 161-177.

Papers in proceedings of international conferences

1. **Sauwen, N.**, Sima, D.M., Acou, M., Achten, E., Maes, F., Himmelreich, U. and Van Huffel, S. A semi-automated segmentation framework for MRI based brain tumor segmentation using regularized nonnegative matrix factorization. International Conference on Signal Image Technology & Internet Based Systems 2016, Naples, Dec. 2016 (Accepted).

2. Bharath, H.N., **Sauwen, N.**, Sima, D.M., Himmelreich, U., De Lathauwer, L., and Van Huffel, S. Canonical polyadic decomposition for tissue type differentiation using multi-parametric MRI in high-grade gliomas. European Signal Processing Conference (EUSIPCO) 2016, Budapest, Aug. 2016 (pp. 547-551)

3. **Sauwen, N.**, Acou, M., Bharath, H.N., Sima, D.M., Veraart, J., Maes, F., Himmelreich, U., Achten, E., and Van Huffel, S. Initializing nonnegative matrix factorization using the successive projection algorithm for multi-parametric medical image segmentation. 24th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), Bruges, May 2016 (pp.265-270).

4. Bharath, H.N., Sima, D.M., **Sauwen, N.**, Himmelreich, U., De Lathauwer, L., and Van Huffel, S. Tensor based tumor tissue type differentiation using magnetic resonance spectroscopic imaging. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Aug. 2015 (pp. 7003-6).

Abstracts in proceedings of international conferences

1. Bharath, H.N., **Sauwen, N.**, Sima, D.M., Himmelreich, U., De Lathauwer, L., and Van Huffel, S. Canonical polyadic decomposition for tissue type differentiation using multi-parametric MRI in high-grade gliomas. 33rd Annual Conference of the European Society of Magnetic Resonance in Medicine and

Biology (ESMRMB 2016), Vienna, Austria, Sept. 2016.

2. Bharath, H.N., **Sauwen, N.**, Sima, D.M., Himmelreich, U., and Van Huffel, S. Non-Negative Canonical Polyadic Decomposition Based Brain Tumor Characterization From Multi-Parametric MR Imaging. TDA 2016 workshop on Tensor Decompositions and Applications, Leuven, Belgium, Jan. 2016.

3. **Sauwen, N.**, Sima, D., Van Cauter, S., Veraart, J., Leemans, A., Himmelreich, U., and Van Huffel, S. Hierarchical non-negative matrix factorization using multi-parametric MRI to assess tumor heterogeneity within gliomas. 23rd Annual Conference of the International Society for Magnetic Resonance in Medicine (ISMRM 2015), Toronto, Canada, Jun. 2015.

4. **Sauwen, N.**, Sima, D.M., Van Cauter, S., Veraart, J., Leemans, A., Himmelreich, U., Van Huffel, S. Hierarchical non-negative matrix factorization using multi-parametric MRI to assess tumor heterogeneity within gliomas. 7th Annual meeting of the International Society for Magnetic Resonance in Medicine, Benelux Chapter (ISMRMB 2015), Ghent, Belgium, Jan. 2015.

5. **Sauwen, N.**, Sima, D., Van Huffel, S., Biopsy targeting based on perfusion and diffusion MR related parameters, IEEE International Symposium on Biomedical Imaging (ISBI 2014), Beijing, China, May 2014.

6. **Sauwen, N.**, Van Cauter, S., Veraart, J., Sima, D.M., Himmelreich, U., Van Huffel, S. Biopsy targeting based on perfusion and diffusion MR related parameters. 6th Annual meeting of the International Society for Magnetic Resonance in Medicine, Benelux Chapter (ISMRMB 2014), Maastricht, the Netherlands, Jan. 2014.

7. **Sauwen, N.P.L.**, Sima, D.M., Van Cauter, S., Veraart, J., Himmelreich, U., Van Huffel, S., Increasing tumor grade separability by combining MR parameters, 30th Annual Conference of the European Society for Magnetic Resonance in Medicine and Biology (ESMRMB 2013), Toulouse, France, Oct. 2013

8. Li, Y. **Sauwen, N.**, Sima, D.M., Croitor Sava, A.R., Van Cauter, S., Himmelreich, U., Pi, Y., Van Huffel, S. Unsupervised tissue type differentiation for glioma diagnosis using Magnetic Resonance Spectroscopic Imaging. 5th Annual meeting of the International Society for Magnetic Resonance in Medicine, Benelux Chapter (ISMRMB 2013), Rotterdam, the Netherlands, Jan. 2013.



FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF ELECTRICAL ENGINEERING
STADIUS DIVISION
Kasteelpark Arenberg 10 box 2446
B-3001 Leuven
nicolas.sauwen@kuleuven.be
http://www.esat.kuleuven.be/stadius