Spring 2017

Improved interpretation of brain anatomical structures in magnetic resonance imaging using information from multiple image modalities

Ali Ghayoor
University of Iowa

IMPROVED INTERPRETATION OF BRAIN ANATOMICAL STRUCTURES IN MAGNETIC RESONANCE IMAGING USING INFORMATION FROM MULTIPLE IMAGE MODALITIES

by

Ali Ghayoor

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Electrical and Computer Engineering
in the Graduate College of
The University of Iowa

May 2017

Thesis Supervisor: Associate Professor Hans J. Johnson

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Ali Ghayoor

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Electrical and Computer Engineering at the May 2017 graduation.

Thesis Committee: _____
Hans J. Johnson, Thesis Supervisor

_____
Gary E. Christensen

_____
Mona K. Garvin

_____
Mathews Jacob

_____
Vincent A. Magnotta

_____
Jatin G. Vaidya

To my parents, Mohtaram and Mostafa, for their unconditional love, support and encouragement through my life to pursue my educational dreams.

To my sister, Maryam, my brother-in-law, Mehrdad, and my brilliant nephews, Parsa and Arya, who were with me during this academic journey and gave me hope and encouragement every single day to keep me going.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere appreciation to my research adviser Dr. Hans Johnson for all his inspiration, patience, and support throughout the difficult and challenging process of research and writing necessary to complete my PhD journey. He showed me how to be enthusiastic and stay positive even during the tough times and how to think differently to figure out the crux of the issue. Without his guidance and support, this journey would have never ended like this. I would also like to thank my dissertation committee members, Drs. Christensen, Garvin, Jacob, Magnotta, and Vaidya for their valuable time, support and insightful suggestions. I learned a lot from many professors in the ECE department. I appreciate all of them especially Dr. Milan Sonka and Dr. Mathews Jacob that the knowledge I gained under their instructions made the basis of this work. I am also grateful for the help of the former and current members of the SINAPSE lab, Dave, Kent, Joy, Regina, Jessica, David, Hui, Mark, Chen, Eric, Kathy, Jacquie, Daniel and Jason. I want to thank PREDICT-HD study group for this opportunity to learn and research.

I would like to extend my sincere gratitude to all my friends in Iowa City. I want to thank my former roommates, Mehrdad Farahani, Alireza Firoozfar, and also Iman Mallakpour for all the years sharing the journey. I especially want to thank Anna Thomas who has rewarded me with endless support, understanding and caring. Thank you all for giving me the beautiful and enjoyable memories that I am going to cherish for a lifetime.

Last, but not least, I would like to thank my family, my father, Mostafa, my mother, Mohtaram, my sisters, Maryam and Elham, my brother-in-laws, Mehrdad and Saeid, my brother, Ehsan, my brilliant nephew Parsa, and my lovely little nephew, Arya, who joined our family during my PhD journey. I do not have enough words to thank you for your endless support and unconditional love. I want to share this honor with you and bless all you have done for me. You are in my heart now and forever.

ABSTRACT

This work explores if combining information from multiple Magnetic Resonance Imaging (MRI) modalities provides improved interpretation of brain biological architecture as each MR modality can reveal different characteristics of underlying anatomical structures. Structural MRI provides a means for high-resolution quantitative study of brain morphometry. Diffusion-weighted MR imaging (DWI) allows for low-resolution modeling of diffusivity properties of water molecules.

Structural and diffusion-weighted MRI modalities are commonly used for monitoring the biological architecture of the brain in normal development or neurodegenerative disease processes. Structural MRI provides an overall map of brain tissue organization that is useful for identifying distinct anatomical boundaries that define gross organization of the brain. DWI models provide a reflection of the micro-structure of white matter (WM), thereby providing insightful information for measuring localized tissue properties or for generating maps of brain connectivity. Multispectral information from different structural MR modalities can lead to better delineation of anatomical boundaries, but careful considerations should be taken to deal with increased partial volume effects (PVE) when input modalities are provided in different spatial resolutions. Interpretation of diffusion-weighted MRI is strongly limited by its relatively low spatial resolution. PVE's are an inherent consequence of the limited spatial resolution in low-resolution images like DWI.

This work develops novel methods to enhance tissue classification by address-

ing challenges of partial volume effects encountered from multi-modal data that are provided in different spatial resolutions. Additionally, this project addresses PVE in low-resolution DWI scans by introducing a novel super-resolution reconstruction approach that uses prior information from multi-modal structural MR images provided in higher spatial resolution.

The major contributions of this work include: 1) Enhancing multi-modal tissue classification by addressing increased PVE when multispectral information come from different spatial resolutions. A novel method was introduced to find pure spatial samples that are not affected by partial volume composition. Once detecting pure samples, we can safely integrate multi-modal information in training/initialization of the classifier for an enhanced segmentation quality. Our method operates in physical spatial domain and is not limited by the constraints of voxel lattice spaces of different input modalities. 2) Enhancing the spatial resolution of DWI scans by introducing a novel method for super-resolution reconstruction of diffusion-weighted imaging data using high biological-resolution information provided by structural MRI data such that the voxel values at tissue boundaries of the reconstructed DWI image will be in agreement with the actual anatomical definitions of morphological data.

We used 2D phantom data and 3D simulated multi-modal MR scans for quantitative evaluation of introduced tissue classification approach. The phantom study result demonstrates that the segmentation error rate is reduced when training samples were selected only from the pure samples. Quantitative results using Dice index from 3D simulated MR scans proves that the multi-modal segmentation quality with low-

resolution second modality can approach the accuracy of high-resolution multi-modal segmentation when pure samples are incorporated in the training of classifier. We used high-resolution DWI from Human Connectome Project (HCP) as a gold standard for super-resolution reconstruction evaluation to measure the effectiveness of our method to recover high-resolution extrapolations from low-resolution DWI data using three evaluation approaches consisting of brain tractography, rotationally invariant scalars and tensor properties. Our validation demonstrates a significant improvement in the performance of developed approach in providing accurate assessment of brain connectivity and recovering the high-resolution rotationally invariant scalars (RIS) and tensor property measurements when our approach was compared with two common methods in the literature.

The novel methods of this work provide important improvements in tools that assist with improving interpretation of brain biological architecture. We demonstrate an increased sensitivity for volumetric and diffusion measures commonly used in clinical trials to advance our understanding of both normal development and disease induced degeneration. The improved sensitivity may lead to a substantial decrease in the necessary sample size required to demonstrate statistical significance and thereby may reduce the cost of future studies or may allow more clinical and observational trials to be performed in parallel.

PUBLIC ABSTRACT

This work explores if combining information from multiple Magnetic Resonance Imaging (MRI) modalities provides improved interpretation of brain biological architecture. Structural MRI and diffusion-weighted MR imaging (DWI) are commonly used modalities for monitoring the biological architecture of the brain in normal development or neurodegenerative disease processes. This work emphasizes on the issue of increased partial volume effects (PVE) when multispectral information from different structural MR modalities are used for tissue classification of brain anatomical structures. Novel methods are developed that lead to better delineation of anatomical boundaries by addressing challenges of PVE algorithmically. Additionally, this study addresses PVE in low-resolution DWI by introducing a novel super-resolution reconstruction approach that uses anatomical priors from high biological-resolution information provided by structural MRI data.

The novel methods of this work provide important improvements in tools that assist with improving interpretation of brain biological architecture. We demonstrate an increased sensitivity for volumetric and diffusion measures commonly used in clinical trials to advance our understanding of both normal development and disease induced degeneration. The improved sensitivity may lead to a substantial decrease in the necessary sample size required to demonstrate statistical significance and thereby may reduce the cost of future studies or may allow more clinical and observational trials to be performed in parallel.

TABLE OF CONTENTS

LIST OF TABLES xii

LIST OF FIGURES xiii

CHAPTER

1 SIGNIFICANCE AND BACKGROUND 1

1.1 Introduction 1
1.1.1 Thesis Aims 4
1.1.2 Thesis Overview 7
1.2 Structural MRI Principles 8
1.3 Diffusion-Weighted Imaging 12
1.4 Diffusion Tensor Imaging 16
1.4.1 Tensor Estimation 17
1.4.2 Correct Errors in Estimation of Diffusion Tensors 20
1.4.3 Calculation of Rotationally Invariant Scalars (RISs) . . . 24
1.4.4 Units of Rotationally Invariant Scalars 26
1.4.5 Distance Metrics for Diffusion Tensors 27
1.4.5.1 L^2 Distance or Frobenius Distance 29
1.4.5.2 Riemannian Geometric-Based Distance 29
1.4.5.3 Kullback-Leibler Distance 30
1.4.5.4 Comparison of Tensor Distance Metrics 31
1.5 Partial Volume Effect (PVE) 31
1.6 PVE Challenges in Processing of Large-scale Heterogeneous Data 36

2 TISSUE CLASSIFICATION OF LARGE-SCALE MULTI-SITE HETEROGENEOUS MR DATA USING FUZZY K-NEAREST NEIGHBOR METHOD 38

2.1 Introduction 38
2.2 Methods 39
2.2.1 General Framework 39
2.2.2 New Classifier 42
2.2.2.1 Training Sample Set 43
2.2.2.2 Test Sample Set 44
2.2.2.3 Run the Algorithm 44
2.3 Experimental Methods 46
2.3.1 Test Data 46
2.3.2 Qualitative Evaluation 46

2.3.3 Quantitative Evaluation 47
2.4 Results 48
2.5 Discussion and Conclusions 48

3 ENHANCE MR MULTI-MODAL TISSUE CLASSIFICATION BY ADDRESSING PARTIAL VOLUME EFFECT 54

3.1 Enhancement of Classification Framework to Run in Physical Space 55
3.1.1 Evaluation 55
3.1.1.1 Test Data 55
3.1.1.2 Results 55
3.1.2 Conclusion 55
3.2 Significance of Multi-modal Classification 56
3.3 Multi-modal Partial Volume Effect 62
3.4 Methods 64
3.4.1 Computing Pure Plugs Mask 64
3.4.1.1 Avoid Inherent PVE 64
3.4.1.2 Avoid Multi-modal PVE 66
3.4.1.2.1 What Is a Pure Plug? 67
3.4.1.2.2 Standardize Intensity 68
3.4.1.2.3 Pure Plugs Integrity Metric 68
3.5 Experimental Methods and Results 75
3.5.1 2D Phantom Analysis 75
3.5.2 Qualitative Evaluation 79
3.5.3 Quantitative Evaluation 79
3.6 Discussion and Conclusions 84

4 SUPER-RESOLUTION RECONSTRUCTION OF LOW-RESOLUTION DIFFUSION-WEIGHTED IMAGING DATA USING *A PRIORI* KNOWLEDGE OF ANATOMICAL STRUCTURES 88

4.1 Introduction 88
4.2 Mathematical Background 91
4.2.1 Regularized Recovery of Inverse Problems 91
4.2.1.1 1-D Computation on Step Edges 93
4.2.2 l_1-norm Minimization in Sparse Signal Recovery 94
4.2.3 Enhance Sparsity Recovery in l_1-norm Minimization . . . 95
4.2.4 Analytical Justification 97
4.2.5 Variation of Weights 100
4.3 Methods 100
4.3.1 Weighted Total Variation Minimization for Image Reconstruction 100
4.3.2 Construction of Weights from Estimated Anatomical Edge Map 102

4.4 Preliminary Assessments on 2D Data 105
4.4.1 Test Data 105
4.4.2 Evaluation Metric 106
4.4.3 Preliminary 2D Results 107
4.4.4 Conclusion 108
4.5 Experimental Methods and Results 112
4.5.1 Test Data 112
4.5.2 **Approach 1**: Evaluation Based on the Overlap Between Tractography Results 113
4.5.3 **Approach 2**: Evaluation Based on the Difference Between Rotationally Invariant Scalar Measurements 118
4.5.4 **Approach 3**: Evaluation Based on the Differences Between Tensor Proprieties 131
4.6 Discussion and Conclusions 136

5 CONCLUDING REMARKS 142

APPENDIX 146

REFERENCES 148

LIST OF TABLES

Table

1.1 Summary of commonly used DTI scalars. 24

2.1 Atlas definition of 15 region-specific intensity-context priors. Each tissue type is sub-divided into regions of interest with given names. (Gm = Gray matter, Wm = white matter, Csf = cerebrospinal fluid, Crbl = Cerebellum, Vb = venous blood). 47

3.1 Test dataset to create a test pure plugs mask. 74

4.1 Test dataset for initial assessments. 105

4.2 The average Bhattacharyya Coefficient to quantify overlap between tracts of interest extracted from the high-resolution baseline image and each of reconstructed images. 118

4.3 Error Mean Value of DTI scalars within each region of interest across all subjects. 119

4.4 Error mean value of all tensor metrics within each region of interest across all subjects. 132

LIST OF FIGURES

Figure

1.1 Hydrogen protons are oriented randomly and their magnetic fields cancel out when there is no external magnetic field. However, when a strong external magnetic field (B_0) is applied, protons are aligned with B_0 field and a net magnetization will be produced parallel to the main magnetic field also called longitudinal direction (Pooley, 2005). 9

1.2 (Left) prior the RF pulse, the net magnetization is parallel to the longitudinal direction. (Center and Right) When RF pulse is applied at Larmor frequency, its energy is absorbed by the protons that rotates the net magnetization from z axis causing a decrease in longitudinal direction and an increase in transversal direction. This figures shows a 90° RF pulse (Pooley, 2005). 11

1.3 (a) T1 relaxation is defined as the time that it takes the longitudinal magnetization to grow back to 63% of its final value. Different tissues have different rates of T1 relaxation. A T1-weighted image is obtained at a time when the T1 relaxation curves are widely separated. (b) T2 relaxation is defined as the time that it takes the transverse magnetization to decrease to 37% of its starting value. Different tissues have different rates of T2 relaxation. T2-weighted contrast will be maximized when the image is obtained at a time when the T2 relaxation curves are widely separated (Pooley, 2005). T1 and T2 contrasts provide complementary intensity profiles for a certain tissue. 13

1.4 The basic diffusion-weighted pulse sequences originated from the Stejskal-Tanner pulsed gradient spin echo technique. The paired diffusion-sensitizing gradients does not affect stationary spins, but diffusing spins are dephased by moving into different locations between the first and second gradients. *"Courtesy Allen D. Elster, MRIquestions.com"*. 14

1.5 The probability of obtaining negative eigenvalues for the 'rice' shaped diffusion tensor as a function of λ_1/λ_3 at different noise levels (Skare et al., 2000). 21

1.6 Comparison of tensor estimation with and without correction methods. Top left: results without correction. Top right: absolute value of eigenvalues are used. Bottom left: negative eigenvalues are reset to zero. Bottom right: nearest symmetric positive semidefinite matrix is found for each invalid tensor, which shows visually cleaner results. 23

1.7 A schematic explanation of the partial volume effect in the context of brain magnetic resonance imaging (Tohka, 2014). 32

2.1 A rough demonstration of iterative classification process by (Kim and Johnson, 2013) using expectation maximization (EM) method. In our development, the EM output posterior probability maps and the input multi-modal MR images are used to train a fuzzy k-NN classifier. 41

2.2 Visual comparison of segmentation results on a sample PREDICT-HD MR data between (a) EM only based classification and (b) developed enhancements using a fuzzy $k-NN$ classifier. More accurate delineation is achieved by the developed method. 49

2.3 Comparison of classification of cerebrospinal fluid (CSF), Grey matter (GM) and White matter (WM) tissues between EM only based classification (black) and the extended method by a fuzzy $k-NN$ classifier (blue) using two independent measures, Dice index (larger is better) and average Hausdorff distance (smaller is better) . The evaluation is performed along three degrees of bias-field (rf=0, rf=20 and rf=40) and six levels of noise (0%, 1%, 3%, 5%, 7%, and 9%) along x-axis. Both similarity measures show improvement on the results of developed $k-NN$ enhancements. . 50

2.4 Two independent measures, Dice index (larger is better) and average Hausdorff distance (smaller is better), are used to compare the classification of cerebrospinal fluid (CSF), Grey matter (GM) and White matter (WM) tissues in two cases: (1) $k-NN$ classifier was trained by thresholded EM tissue probability maps (equation 2.8) (results in blue); (2) $k-NN$ classifier was trained directly by EM tissue probability maps (results in red). The evaluation is performed along three degrees of bias-field (rf=0, rf=20 and rf=40) and six levels of noise (0%, 1%, 3%, 5%, 7%, and 9%) along x-axis. Both similarity measures show improvement on using the thresholded EM tissue probability maps in the training phase of $k-NN$. 53

3.1 The classification framework generates essentially the same results in single-modal mode when it is run in physical space versus voxel space. Blue lines are masked by the red lines as the results are overlapped. 57

3.2 (a) It is hard to distinguish between gray matter (GM) and venous blood (VB) (as both are shown as grey) and to distinguish CSF from the bone marrow (since both are shown as dark) by looking only at T1-weighted MR scan. (b) T2-weighted MR provides better contrast in the case of T1 ambiguities between mentioned tissue types, since here GM is gray but VB is dark; also CSF is shown as white while bone marrow is dark. . . . 58

3.3 Visual comparison of segmentation results on a sample PREDICT-HD MR dataset (a) Single-modal segmentation of CSF using only T1-weighted modality. Segmentation results are shown in blue boundaries while the true anatomical boundary is defined in yellow based on the contrast provided by T2-weighted modality. (b) Multi-modal segmentation of CSF using both T1/T2-wegihted scans shows enhanced delineation of CSF as the segmentation results conform with true anatomical boundaries. . . . 59

3.4 Blue lines show less accurate and less robust classification results when second modality ($T2$: $1 \times 1 \times 3\ mm^3$) has lower resolution than the first modality ($T1$: $1 \times 1 \times 1\ mm^3$). 61

3.5 $1D$ demonstration of partial volume effect (PVE) when MRI signals of different modalities are sampled to discredited segments with different spatial resolutions and origins. Question mark represent the sample values that reflect a partial volume composition of more than one tissue type. . 65

3.6 (a) a coronal view of a scatter point map of two Gray matter (black) and White matter (pink) tissues. (b) only pure samples are shown. (c) only mixed samples are shown. 66

3.7 Anatomical tissue boundaries are detected using a Canny edge detector. Detected edges should be excluded from the pure plugs mask to avoid inherent PVE in tissue boundaries. 67

3.8 Intensity transform function applied to each input modality image to make all the input images have the same intensity dynamic range. 69

3.9 Joint image histogram for two modality images. Volume 1 is a T1-weighted image, and volume 2 is a T2-weighted scan from the same subject. The distribution of points is shown in different colors for background (grey) and four tissue regions: White Matter (red), Gray Matter (blue), Cerebrospinal Fluid (green), and Venous Blood (yellow). A box is plotted around the mean of each region with the size of 4 standard deviation at each direction. 70

3.10 Mahalanobis distance is scale invariant and does not consider the range of scattered points. (a) The Mahalanobis distance for all points is 1.22, and all points belong to a single tissue region. (b) The Mahalanobis distance for all points is 1.22, but points belong to four different tissue regions. . . 72

3.11 A test pure plugs mask created from 3 modality scans from a subject test listed in table (3.1). 75

3.12 Two 2D phantom images representing two modalities and their corresponding baseline label map. 77

3.13 (a) The second modality downsampled by a factor of 10. (b) The corresponding pure plugs mask (in blue) overlaid on the downsampled image. 78

3.14 The segmentation error rate is computed as the percentage of misclassified pixels to the total number of pixels when the baseline label map is compared with the results of multi-modal segmentation from three different experiments: (1) The low-resolution modality was upsampled using a nearest neighbor interpolation. (2) A fuzzy K-Nearest Neighbor classification was run when a random set of samples were used to train the classifier. (3) A fuzzy K-Nearest Neighbor classification was run when only pure samples were used to train the classifier. 80

3.15 Training samples and their corresponding class labels in the feature space. (a) Only pure samples are using for training. (b) Both pure and non-pure samples are used for training. 81

3.16 (a) Segmentation results when pure plugs mask is not incorporated in classification process. (b) Results when only pure samples are used to initialize/train the classification methods. Incorporating pure samples in the classification process enhances segmentation in tissue boundaries and can help better delineation of subtle tissue regions. 82

3.17 Green lines show the results when the multi-modal segmentation is run using two modalities with the same isotropic 1 mm^3 resolution. Blue lines show the results when $T2$ modality has a lower spatial resolution as $1 \times 1 \times 3$ mm^3. Dashed lines show the improved results when only pure samples are involved in the classification process. This improvement is more significant when the second modality has lower spatial resolution as indicated by blue lines. 83

3.18 Human brain diagram showing four different lobes of brain *(Image: Public domain)*. 84

3.19 Regional benefits of using pure samples in segmentation of gray matter (GM) in in each of four different lobes of brain (frontal, occipital, temporal and parietal). Green lines show the results when the multi-modal segmentation is run using two modalities with the same isotropic 1 mm^3 resolution. Blue lines show the results when $T2$ modality has a lower spatial resolution as $1 \times 1 \times 3$ mm^3. Dashed lines show the improved results when only pure samples are involved in the classification process. 85

3.20 Regional benefits of using pure samples in segmentation of white matter (WM) in in each of four different lobes of brain (frontal, occipital, temporal and parietal). Green lines show the results when the multi-modal segmentation is run using two modalities with the same isotropic 1 mm^3 resolution. Blue lines show the results when $T2$ modality has a lower spatial resolution as $1 \times 1 \times 3$ mm^3. Dashed lines show the improved results when only pure samples are involved in the classification process. 86

4.1 The log-sum concave penalty function $f_{log,\epsilon}(t)$ is a better approximation for the l_0 sparsity count $f_0(t)$ rather than the traditional convex l_1 regularization $f_1(t)$ (Candes *et al.*, 2008). 97

4.2 Intensity transform function. Strong edges (with values above %95 percentile) are mapped to a maximum value M. Weak edges (with values below %50 percentile) are mapped to a value $\epsilon > 0$ close to zero. Other edges are mapped linearly to a $\epsilon < value < M$. 104

4.3 Created spatial weights based on the anatomical edges estimated from the high-resolution structural MR modalities. (a) T1-weighted MR image. (b) T2-weighted MR image. (c) Estimated edge map based on the maximum gradient values as described in section 4.3.2. (d) Prior spatial weights created from estimated edge map based on equation (4.36). 107

4.4 Reconstructed b_0 component of a typical DWI subject using 3 different methods. The SNR value is provided for the result of each method demonstrated in the first row. The second row shows the low-resolution image, by a factor of 2, and the difference image between each reconstructed image and the original high-resolution baseline image, that is magnified by 20 times. The higher SNR value for the developed weighted-TV approach indicates better reconstruction results. 109

4.5 Reconstructed first gradient component of a typical DWI subject using 3 different methods. The SNR value is provided for the result of each method demonstrated in the first row. The second row shows the low-resolution image, by a factor of 2, and the difference image between each reconstructed image and the original high-resolution baseline image, that is magnified by 20 times. The higher SNR value for the developed weighted-TV approach indicates better reconstruction results for the gradient components as well. 110

4.6 Reconstructed b0 component of a typical DWI subject using FRI edge map suggest in (Ongie and Jacob, 2015). 111

4.7 (Left): Arcuate Fascicle (AF) from right view. (Right): Cortico-spinal tract (CST) from anterior view. 114

4.8 A sample computation of Bhattacharyya coefficient based on the probability distribution function (pdf) on each of x, y and z coordinates. In this example BC is calculated as 0.98. 116

4.9 The Bhattacharryya coefficients computed for 20 test datasets to quantify the overlap for cortico-spinal tract (CST) and arcuate fascicle (AF) extracted from the HR baseline and each of reconstructed images by three different methods: The developed WTV, Standard TV, and zero-padded IFFT. 117

4.10 Comparing the average fractional anisotropy (FA) values within 4 WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT (IFFT). The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well. 120

4.11 Comparing the average mean diffusivity (MD) values within 4 WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT (IFFT). The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well. 121

4.12 Comparing the average radial diffusivity (RD) values within four WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT (IFFT). The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well. 122

4.13 Comparing the average of axial diffusivity (AD) values within 4 WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT (IFFT). The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well. 123

4.14 Scatter plots of $\Delta error$ of FA in each WM region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit. 127

4.15 Scatter plots of $\Delta error$ of MD in each WM region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit. 128

4.16 Scatter plots of $\Delta error$ of RD in each WM region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit. 129

4.17 Scatter plots of $\Delta error$ of AD in each WM region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit. 130

4.18 Comparing the average Riemannian error values within four WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT. The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well. 133

4.19 Comparing the average Kullback-Leibler error values within four WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT. The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well. 134

4.20 Comparing the average Frobenius norm values within four WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT. The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well. 135

4.21 Scatter plots of $\Delta error$ of Riemannian distance in each white matter (WM) region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit. 137

4.22 Scatter plots of $\Delta error$ of Kullback-Leibler distance in each white matter (WM) region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit. 138

4.23 Scatter plots of $\Delta error$ of Frobenius norm in each white matter (WM) region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit. 139

CHAPTER 1
SIGNIFICANCE AND BACKGROUND

1.1 Introduction

This work develops novel methods that incorporate information from multiple imaging modalities for improved interpretation of brain anatomical structures from magnetic resonance imaging (MRI) data. The project is organized around two primary efforts: 1) First, novel methods are developed to enhance multispectral tissue classification by incorporating complementary information from multiple structural MR modality scans with different spatial resolutions. The approach allows information acquired from low-resolution MR scans (i.e. T2-weighted, PD-weighted or FLAIR) to improve segmentation quality over previously produced results that utilized only high-resolution T1-weighted structural MR images by addressing the challenges of partial volume effects that are encountered from multi-modal data. 2) Second, we develop a novel method for super-resolution reconstruction using the prior anatomical information extracted from high spatial resolution images to improve biological interpretation of low-resolution modality sources. For example, we incorporate the high biological resolution of brain anatomy description from structural MR (T1/T2-wighted) scans to improve the super-resolution reconstruction of diffusion-weighted imaging (DWI) data from the same subject. The developed method aims to increase the biological interpretability of input low-resolution DWI to a high-resolution DWI by a factor of 8 (2.0^3 mm^3).

This research project focuses on challenges encountered in processing of a heterogeneous dataset in multi-center neurodegenerative studies like PREDICT-HD [1] study of Huntington's Disease (HD). We describe novel approaches for the segmentation and super-resolution reconstruction (SRR), that help improve interpretability of data and can, for example, increase the sensitivity of volumetric and diffusion measures used by clinicians to improve the clinical and observational trials to test therapies that may slow the progression of disease. Increasing the sensitivity of measures can also reduce the number of samples needed for longitudinal and cross-sectional analysis that can potentially reduce the cost of future studies.

In multi-center studies of neurodegenerative disorders such as Alzheimer's disease, Schizophrenia and Huntington's Disease, different MR modalities are often acquired at each scanning session, since each scan modality reveals different characteristics of underlying biological architecture. There are several types of MR contrasts that result in different MR modality images:

1. Static contrast that is sensitive to relaxation properties of the spins (T1 and T2 relaxations). Several structural MR modalities (T1-weighted, T2-weighted, Proton density (PD), and Flair) are generated based on static contrast [2, 3, 4].
2. Endogenous contrast that depends on intrinsic properties of tissues (e.g. fMRI BOLD) [5].
3. Exogenous contrast that depends on an injected agent contrast like Gadolinium (Gd). An example of MR image modality using this type of contrast is dynamic contrast-enhanced T1-weighted MRI (DCE-MR) [6].

4. Motion contrast that is sensitive to movement of spins through space e.g. diffusion-weighted imaging (DWI) modality that is generated based on diffusion patterns of water molecules in biological tissues under different sensitizing gradient directions [7].

The algorithmic developments presented in this work are focused on structural and diffusion-weighted MR modalities. T1-weighted MR is the most used modality for the segmentation of brain anatomical structures; however, it is hard to distinguish all biological tissue types (e.g. CSF from the bone marrow, or the gray matter from the blood vessels) by looking only at T1 scan. Using complementary information from other modalities (e.g. T2-weighted scan) can provide better contrast in the case of above-mentioned T1-only tissue type ambiguities. However, if T2-weighted scan is provided in lower spatial resolution, it can adversely affect the multi-modal segmentation results as it increases partial volume effects (PVE) issue in tissue boundaries.

Diffusion-weighted imaging (DWI) is another modality that maps water molecule diffusion patterns in biological tissue that can reveal abnormalities in white matter fiber structure and provides models of brain connectivity. Multiple 3D volumes are collected in a single 4D DWI scan, where each 3D volume quantifies diffusion in a different direction. The different sensitizing gradient directions are created by changing the direction of pulsed gradient fields. In a 4D DWI scan, a baseline image or $b0$ image without diffusion weighting is also needed to quantify the diffusion in the DWI data. Diffusion-weighted MR imaging been shown to be sensitive to microscopic degeneration of white matter that is not measurable in structural MRI. Measuring diffusion

properties of functional white matter areas in the time course of degeneration will help us to advance our understanding of neuropathological basis in neurodegenerative diseases. DWI is, however, strongly limited by its relatively low spatial resolution. A DWI voxel volume is approximately 8 times larger than that of a typical structural MRI.

The first part of this research project develops novel methods to use complementary information from multiple modalities to enhance the classification results previously generated from high-resolution structural MRI data alone. The second part uses high biological-resolution information provided by structural MRI data for guiding super-resolution reconstruction of each of the 3D DWI sub-volume gradients to provide more biologically interpretable DWI analysis results.

1.1.1 Thesis Aims

This project is organized around two specific aims:

Aim 1: Enhance tissue classification by incorporating multiple modalities with different spatial resolutions

We show increased sensitivity and interpretability of data by carefully using information from multiple modality channels. We develop a novel method to overcome the challenges of partial volume effects that are encountered from multi-modal data (especially at different resolutions).

The challenges are described and compared against failures that occur when naïvely adding multiple modalities.

- **Subtask 1: Tissue classification of large-scale multi-site heterogeneous MR data using fuzzy k-nearest neighbor method**

  This subtask aims to improve automated classification of brain tissues for multi-center 3D MRI data analysis. Previous studies have developed a robust multi-modal tool for automated registration, bias-field correction and tissue classification based on expectation maximization (EM) method [8] that is group specific and uses *a priori* knowledge for all the subjects in an atlas-based approach. In subtask 1, however, we emphasized the importance of a non-parametric model's utility in neurodegenerative diseases, since each subject has unique anatomical states in longitudinal degenerative studies that may not be represented by prior probability distributions. Enhancements are suggested by augmenting the EM-based classification using a fuzzy k-nearest neighbor (k-NN) classifier that builds up a model for each individual subject and complements the classification results that EM produces. Fuzzy k-NN is a non-parametric algorithm, that means it does not make any assumptions on the underlying data distribution and is not biased by prior probability distributions; therefore, it can complement the results of a Gaussian distribution based mixture method in complex decision boundaries.

- **Subtask 2: Enhance multi-modal classification when complementary information comes from a second modality with lower spatial resolution**

  The second task investigates the partial volume effect (PVE) on multi-modal

classification when one modality channel provides information in lower spatial resolution. Partial volume effect complicates the segmentation process, and, due to the complexity of human brain anatomy, addressing the PVE issue is a critical factor for accurate brain structure classification from multiple modalities. Therefore, a novel method is developed to identify spatial samples that are not affected by partial volume composition, termed as *pure samples*, to train/initialize the classification algorithms.

Aim 2: Super-resolution reconstruction of low-resolution diffusion-weighted imaging data using *a priori* knowledge of anatomical structure

The second aim introduces novel approaches to use the high-resolution representation of anatomical structures provided by the structural MR data as *a priori* knowledge to enhance the resolution (in a biologically relevant manner) of diffusion-weighted imaging (DWI) data.

The use of diffusion-weighted imaging data is limited by its relatively low spatial resolution. The resolution of a typical DWI data is $2\times2\times2\ mm^3$, that is 8 times larger than that in structural MRI. The goal of this section is to increase the spatial resolution of an input DWI scan to match the typical anatomical resolution presented in the isotropic ($1\ mm^3$) voxel of structural MR images. However, using only basic interpolation techniques to resample the input DWI data to a higher resolution lattice space introduces no new information or improvement to interpretation of data that can help further medical imaging analysis. This aim investigates the use of tissue

boundary information from the high-resolution structural MRI data to achieve an edge-guided super-resolution reconstruction of DWI scans based on a weighted total variation (WTV) approach, where the voxel values at tissue boundaries will be in agreement with the actual anatomical definitions of morphological data.

1.1.2 Thesis Overview

This thesis is organized in five chapters as follows:

- The remaining of Chapter 1 provides an overview of the technical background utilized in this work and summarizes the previous works for addressing the challenges caused by partial volume effects.
- Chapter 2 explains the methods and validation performed to enhance the tissue classification of large-scale multi-site heterogeneous MR data by augmenting the EM-based classification using a fuzzy k-nearest neighbor (k-NN) classifier.
- Chapter 3 describes the advantage of using information from multiple modality channels in the segmentation of human brain and explains the challenges of partial volume effects that are encountered from multi-modal data (especially at different resolutions). Finally, it describes the methods and validations performed to address the identified challenges and compares the results against failures that occur when data are naïvely added from multiple modalities.
- Chapter 4 provides information regarding the methods and validations performed for an edge-guided weighted total variation (WTV) super-resolution reconstruction approach to increase the spatial resolution of diffusion-weighted

imaging data using *a priori* knowledge of anatomical structures derived from corresponding structural MR data acquired in higher spatial resolution.

- Chapter 5 concludes the remarks of this work.

1.2 Structural MRI Principles

Magnetic Resonance Imaging (MRI) is a medical imaging technique that uses magnetic fields, radio waves, and field gradients to image the anatomy and the physiological processes of the body. This section is an overview on the basic concepts of MR imaging provided in [2, 3, 4]. Magnetic resonance (MR) phenomenon was discovered by Felix Bloch and Edward Purcell independently in 1946. The nuclear magnetic resonance imaging is based on the concept that the nuclei of some atoms absorb and re-emit radio frequency when placed in a magnetic field.

MRI techniques for human scanning are almost exclusive to measuring hydrogen atoms associated with fat and water molecules. Protons in the nuclei of hydrogen atoms are positively charged and spin about their axis, so they acts like tiny magnets in our body. These tiny magnets are randomly oriented, so their magnetic fields cancel out. However, when we place these protons in a strong magnetic field, they will tend to align in parallel or anti-parallel to the magnetic field. While the magnetic fields from many protons will cancel out, a slight excess of the protons will produce a net magnetization parallel to the main magnetic field that is also called the longitudinal direction. This net magnetization becomes the source of our MR signal [3]. This is illustrated in Figure (1.1).

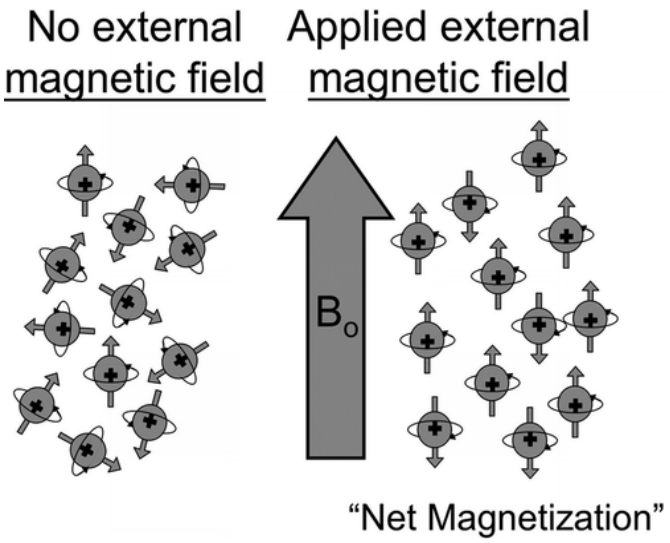


Figure 1.1: Hydrogen protons are oriented randomly and their magnetic fields cancel out when there is no external magnetic field. However, when a strong external magnetic field (B_0) is applied, protons are aligned with B_0 field and a net magnetization will be produced parallel to the main magnetic field also called longitudinal direction (Pooley, 2005).

The force from the magnetic field interacts with the spinning protons and results in precession of the protons. This precessional frequency is determined from the Larmor equation:

$$f = \gamma \times B_0 \tag{1.1}$$

Where f is the frequency of precession, B_0 is the main magnetic field strength, and γ is gyromagnetic ratio that is a characteristic of each type of nuclei and is equal to 42.6 MHz/T (Megahertz per Tesla) for hydrogen protons.

Applying a radio frequency (RF) pulse at the Larmor frequency of the protons causes resonance and an efficient energy transfer from the RF coil to the protons. A brief exposure of RF pulse causes the net magnetization rotates away from the longitudinal direction, which results in a decrease in the longitudinal magnetization while increases the transversal magnetization (Figure 1.2). The amount of rotation from the longitudinal direction is called as the flip angle and depends on the strength and duration of RF pulse. After the RF pulse, the longitudinal magnetization will grow back over time in a direction parallel to the main magnetic field, while transversal magnetization decreases as it begins to dephase.

The time that it takes the longitudinal magnetization to grow back to 63% of its final value is described by a time constant called longitudinal relaxation or $T1$ relaxation and is a characteristic of tissue. The different rate of regrowth of longitudinal magnetization for different tissues is the fundamental source of contrast in T1-weighted images (Figure (1.3 a)). White matter has a very short T1 time and relaxes rapidly. Cerebrospinal fluid (CSF) has a long T1 and relaxes slowly. Gray

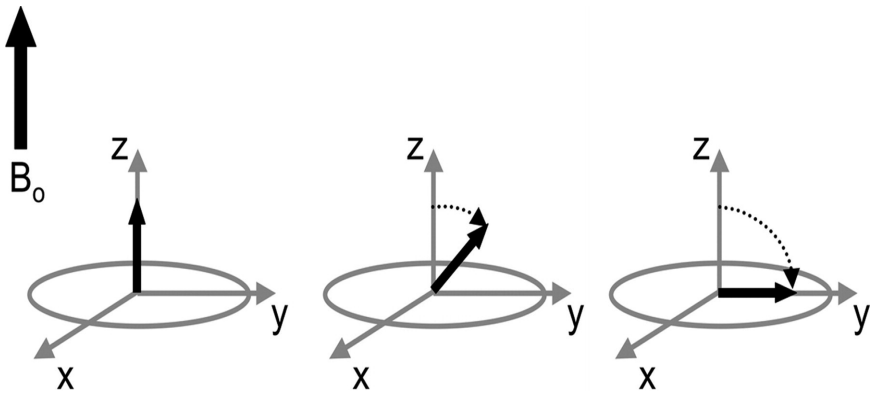


Figure 1.2: (Left) prior the RF pulse, the net magnetization is parallel to the longitudinal direction. (Center and Right) When RF pulse is applied at Larmor frequency, its energy is absorbed by the protons that rotates the net magnetization from z axis causing a decrease in longitudinal direction and an increase in transversal direction. This figures shows a 90° RF pulse (Pooley, 2005).

matter has an intermediate T1 and relaxes at an intermediate rate.

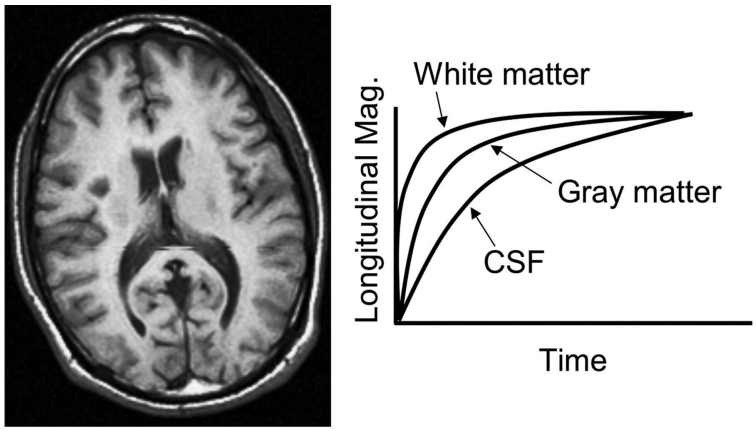
Also, the time that it takes the transverse magnetization to decrease to 37% of its starting value is described by a time constant called transversal relaxation or T2 relaxation that provides another contrast mechanism in MR images. Different tissues have different values of T2 and dephase at different rates (Figure (1.3 b)). White matter has a short T2 and dephases rapidly. CSF has a long T2 and dephases slowly. Gray matter has an intermediate T2 and dephases intermediately. T1-weighted and T2-weighted contrasts present distinguished characteristics for an identical tissue types as shown in Figure (1.3). This complimentary intensity profile between T1- and T2-weighted MRI for a certain tissue type often allows a better

delineation of tissue types.

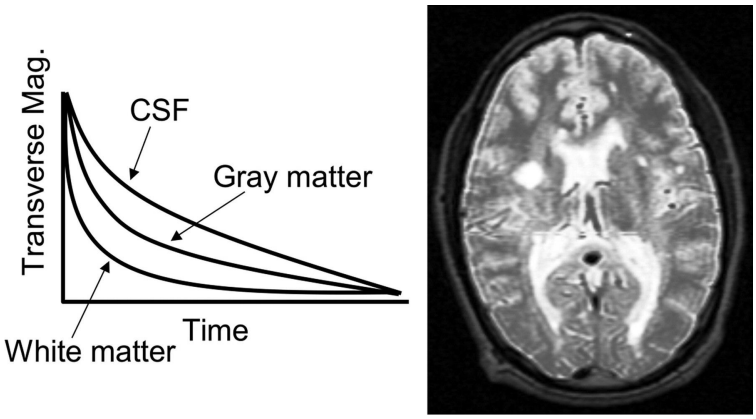
To create different contrasts, there are several MR pulse sequences to control the timing of certain events during MR acquisition. The events include RF pulses and the signal that is formed from these pulses. Other events in a pulse sequence are related to the time and duration of gradient pulses. Gradients are applied to alter the main magnetic field in a predictable fashion and are responsible for localizing the signals from protons located at different positions. Frequency encoding and Phase encoding gradients are applied for individual voxel localization. More details on different pulse sequences are provided in [2, 4]. Study the concepts of MR acquisition is important to understand the source of different MR artifacts [9] and enables the design of optimized post-processing algorithms to enhance the MR image qualities.

1.3 Diffusion-Weighted Imaging

Diffusion Weighted Imaging (DWI) is an imaging method based upon measuring the Brownian motion of water molecules to generate contrast in MR images. Diffusion is measured with a pulsed gradient spin echo (PGSE) technique developed by Stejskal and Tanner [10]. As shown by Figure (1.4), this sequence applies symmetric, strong diffusion-sensitizing gradients (DG's) on either side of the 180°-pulse to spatially encode diffusion. The DG pair does not affect the phases of stationary spins since any phase accumulation from the first gradient lobe is reversed by the second one. However, diffusing spins move into different locations between the first and second lobes and loose signal due to falling out of phase, so they will look darker in the



(a) T1-weighted



(b) T2-weighted

Figure 1.3: (a) T1 relaxation is defined as the time that it takes the longitudinal magnetization to grow back to 63% of its final value. Different tissues have different rates of T1 relaxation. A T1-weighted image is obtained at a time when the T1 relaxation curves are widely separated. (b) T2 relaxation is defined as the time that it takes the transverse magnetization to decrease to 37% of its starting value. Different tissues have different rates of T2 relaxation. T2-weighted contrast will be maximized when the image is obtained at a time when the T2 relaxation curves are widely separated (Pooley, 2005). T1 and T2 contrasts provide complementary intensity profiles for a certain tissue.

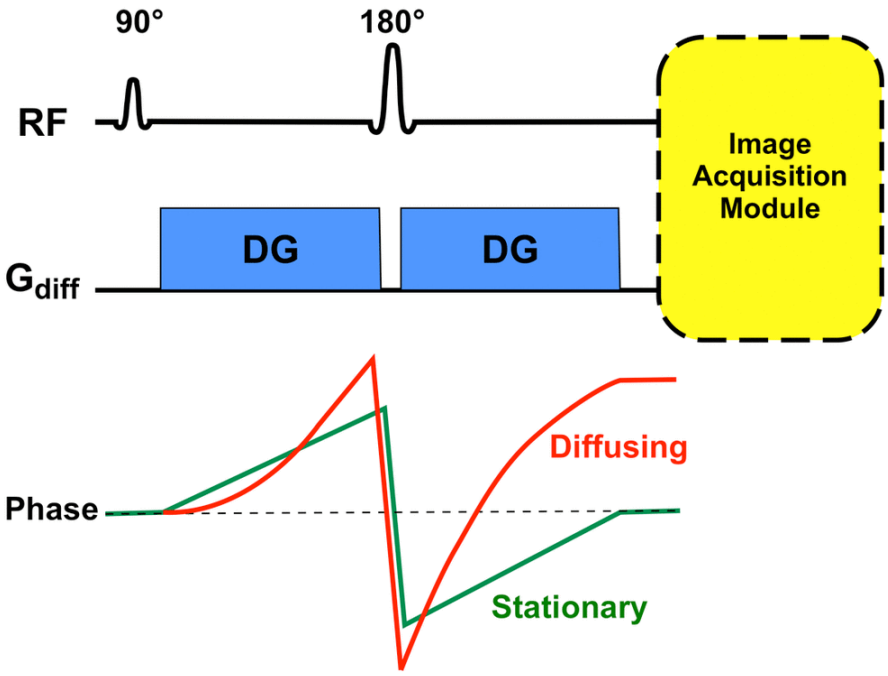


Figure 1.4: The basic diffusion-weighted pulse sequences originated from the Stejskal-Tanner pulsed gradient spin echo technique. The paired diffusion-sensitizing gradients does not affect stationary spins, but diffusing spins are dephased by moving into different locations between the first and second gradients. *"Courtesy Allen D. Elster, MRIquestions.com"*.

output image. The faster the spins move in the direction of the diffusion-sensitizing gradient, the greater the signal loss [10, 11].

Immediately following the second DG, a *rapid* image acquisition module is applied to minimize the effects of bulk motion on the diffusion-weighted images. This module is typically an echo-planar sequence using rapidly oscillating phase and frequency gradients that generate multiple gradient echoes.

To generate 4D diffusion-weighted images (DWI), the DW pulse sequence is first run with the DG's turned off or set to a very low value. A DWI image without a diffusion-sensitizing gradient looks similar to a standard morphometric T2-weighted image because it is basically an echo-planar image (EPI). They are called $b0$ ("b-zero") images and will serve as a baseline for later calculated maps. The DW sequence is then run with the DG's turned on individually or in combination and at various strengths. This produces DW source images sensitized to diffusion in multiple different directions. These gradient scans have areas of signal loss that represents diffusion occurring during the scan.

Signal loss is also dependent on the strength and duration of the diffusion-sensitizing gradient field, also known as the diffusion-weighted factor or b factor (units = sec/mm^2), and D, the apparent diffusion coefficient or ADC (units = mm^2/sec). D is the velocity measure of diffusion. The stronger the diffusion, the greater the apparent diffusion coefficient [12]. The formula for b is:

$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3) \tag{1.2}$$

Where γ is the gyromagnetic ratio, Δ is the time interval between the two diffusion gradient pulses, while G is the magnitude and δ is the duration of each of gradient pulses. If S_0 is the MR signal at baseline, without diffusion-sensitizing gradients, the signal S after applying the diffusion gradients is given by:

$$S = S_0 e^{-bD} \tag{1.3}$$

The value of b is selected by the operator prior to imaging. The choice of b-value

controls the degree of observed diffusion-weighting similar to the way choosing TE affects T2-weighted MRI. Thus, diffusion may be thought of as another relaxation mechanism in addition to T1 and T2. The diffusion becomes the dominant relaxation mechanism of tissue contrast when diffusion gradients are applied.

The b factor is a predetermined parameter for the scanning protocol, S and S_0 are empirically measured by the scanner during each image acquisition, and D is estimated after the scan from equation (1.3) to quantify the diffusion from the imaging information:

$$D = (ln(S_0) - ln(S))/b \tag{1.4}$$

1.4 Diffusion Tensor Imaging

The scalar, D, can also be expressed as a vector, $\mathbf{D}$, that contains the six unique components of the estimated symmetric diffusion tensor. $\mathbf{D}$ can be calculated at each voxel to produce a diffusion tensor image (DTI) that is one of the most popular models for estimating diffusion. Diffusion tensor imaging (DTI) is a special kind of DWI that provides quantitative information with which to visualize and study of white matter connectivity in the brain. Each voxel of DTI contains a 3×3 diffusion tensor that is a symmetric semi-definite matrix modeling the local diffusion within the voxel.

1.4.1 Tensor Estimation

Diffusion Tensor (DT) is the covariance matrix of a Gaussian distribution that models the local diffusion at the voxel level.

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \tag{1.5}$$

Note that $\mathbf{D}$ is a symmetric matrix with six unique components. Diffusion tensors are estimated within each voxel from a 4D DWI data based on Stejkal-Tanner equation, where at least six diffusion scans and one $b0$ image are needed. Consider the Stejkal-Tanner equation:

$$S_i = S_0 e^{-b g_k^T \mathbf{D} g_k} \tag{1.6}$$

The b factor is a predetermined parameter for the scanning protocol. S_0 is the the baseline signal, without diffusion-sensitizing gradients, and S_i is the diffusion signal after applying the diffusion gradients g_i:

$$g_i = [G_x, G_y, G_z]^T \tag{1.7}$$

A simple approach to estimate $\mathbf{D}$ is the "H" approach. We can solve equation (1.6) in terms of $\mathbf{D}$:

$$-ln(\frac{S_i}{S_0})/b = (G_{xi}^2 D_{xx} + G_{yi}^2 D_{yy} + G_{zi}^2 D_{zz} + 2G_{xi}^2 G_{yi} D_{xy} + 2G_{xi}^2 G_{zi} D_{xz} + 2G_{yi}^2 G_{zi} D_{yz}) \tag{1.8}$$

Where the subscript i denotes each gradient direction pulse. Using matrix algebra is one approach to solve this system of equation. Let express $\mathbf{D}$ as a six-element column vector, $\mathbf{d}$ to express the six unique diffusion tensor elements:

$$\mathbf{d} = [D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}]^T \tag{1.9}$$

Each gradient encoding matrix is derived from the normalized gradient components, G_{xi}, G_{yi}, and G_{zi}, (for $i = 1$ to M number of gradients) and is represented as a six-element row matrix, $\mathbf{H}_i$ (Eq. 1.10). The $\mathbf{H}_i$ vectors are combined into a large $M \times 6$ matrix, where M is the number of gradient measurements whose $b > 0$.

$$\mathbf{H}_i = [G_{xi}^2, G_{yi}^2, G_{zi}^2, 2G_{xi}G_{yi}, 2G_{xi}G_{zi}, 2G_{yi}G_{zi}] \tag{1.10}$$

$$\mathbf{H} = \begin{pmatrix} G_{x1}^2 & G_{yi}^2 & G_{zi}^2 & 2G_{xi}G_{yi} & 2G_{xi}G_{zi} & 2G_{yi}G_{zi} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ G_{xM}^2 & G_{yM}^2 & G_{zM}^2 & 2G_{xM}G_{yM} & 2G_{xM}G_{zM} & 2G_{yM}G_{zM} \end{pmatrix} \tag{1.11}$$

Observed data is expressed as the individual measured ADCs. Let's define the Y matrix for the left side of equation (1.8) as:

$$\mathbf{Y}_i = ln\left(\frac{S_0}{S_i}\right)/b \tag{1.12}$$

$$\mathbf{Y} = \left[\frac{ln\left(\frac{S_0}{S_1}\right)}{b}, \frac{ln\left(\frac{S_0}{S_2}\right)}{b}, \cdots, \frac{ln\left(\frac{S_0}{S_M}\right)}{b}\right]^T \tag{1.13}$$

where S_i is the observed signal for the g_i gradient (Eq. 1.7). Note that Y is a volumetric image with noisy data and is expressed as:

$$\mathbf{Y} = \mathbf{Hd} + \boldsymbol{\eta} \tag{1.14}$$

for each acquisition. If $M = 6$, $\mathbf{d}$ can be determined by the following calculation (where $\boldsymbol{\eta} = 0$ since $\boldsymbol{\eta}$ is not able to be distinguished from $\mathbf{d}$) [13]:

$$(\mathbf{H}^{-1}\mathbf{H})\mathbf{d} = \mathbf{d} = \mathbf{H}^{-1}\mathbf{Y} \tag{1.15}$$

However, if $M > 6$, the pseudoinverse of the $\mathbf{H}$ matrix must be used since $\mathbf{H}$ is not a square matrix and thus has no true inverse. The equation (1.14) can still be

solved by first calculating the $\mathbf{H}^{\psi}$ that is the pseudoinverse of $\mathbf{H}$:

$$\mathbf{H}^{\psi}\mathbf{H} = \mathbf{I}_{6\times 6} \tag{1.16}$$

The matrix product $\mathbf{H}^T\mathbf{H}$ can be formed as a square 6×6 matrix and thus has a true inverse. Assumed that $\boldsymbol{\eta} = 0$ in equation (1.14):

$$\mathbf{H}^T\mathbf{Y} = \mathbf{H}^T\mathbf{H}\mathbf{d}$$

$$(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{H}\mathbf{d} = \mathbf{d} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{Y}$$

$$\mathbf{H}^{\psi} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T \tag{1.17}$$

The pseudoinverse can also be obtained through singular value decomposition (SVD). The $\mathbf{H}$ matrix is decomposed into the product of three matrices ($\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$):

$$\mathbf{U} = \text{M} \times 6 \text{ column orthogonal}$$

$$\mathbf{V} = 6 \times 6 \text{ row and column orthogonal}$$

$$\mathbf{W} = 6 \times 6 \text{ diagonal}$$

$$\mathbf{V}\mathbf{V}^T = \mathbf{I}_{6\times 6}$$

$$\mathbf{U}^T\mathbf{U} = \mathbf{I}_{6\times 6}$$

$$(\mathbf{W}^{-1})_{ij} = \frac{1}{\mathbf{W}_{ij}}$$

$$\mathbf{H} = \mathbf{U}\mathbf{W}\mathbf{V}^T \tag{1.18}$$

and the pseudoinverse of $\mathbf{H}$ can be derived by [13]:

$$\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\mathbf{H} = \mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\mathbf{U}\mathbf{W}\mathbf{V}^T = \mathbf{I}_{6\times 6} \tag{1.19}$$

$$\mathbf{H}^{\psi} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T = \mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T \tag{1.20}$$

This section provided the basic concepts to understand the tensor estimation from diffusion-weighted imaging components. More details on other tensor estimation approaches are provided in [14, 13].

1.4.2 Correct Errors in Estimation of Diffusion Tensors

Modeling of water diffusion in the diffusion tensor imaging (DTI) provides a unique noninvasive approach to investigate and characterize the microstructural properties of white matter tissues in brain. Several rotationally invariant scalars (RISs) are calculated to measure important information from the diffusion tensors. However, interpretations and comparisons of these RISs are influenced by the accuracy of the diffusion tensor estimates. The presence of noise in DWI measurements can cause error in estimation of diffusion tensors.

In DTI estimation, diffusion tensors are formally constrained to be in the form of symmetric positive semidefinite square (3×3) matrices. The eigenvalues of a proper symmetric positive semidefinite diffusion tensor $\mathbf{D}$ are always non-negative. Therefore, all scalars derived from a valid tensor assume the use of $\lambda_1 >= \lambda_2 >= \lambda_3 >= 0$ to refer to the eigenvalues of the symmetric positive semidefinite diffusion tensor $\mathbf{D}$.

However, estimated tensors are not guaranteed to be always positive semidefinite due to the presence of noise in the measurement of DWI images. Skare et al. has investigated the probability of obtaining negative eigenvalues as a function of

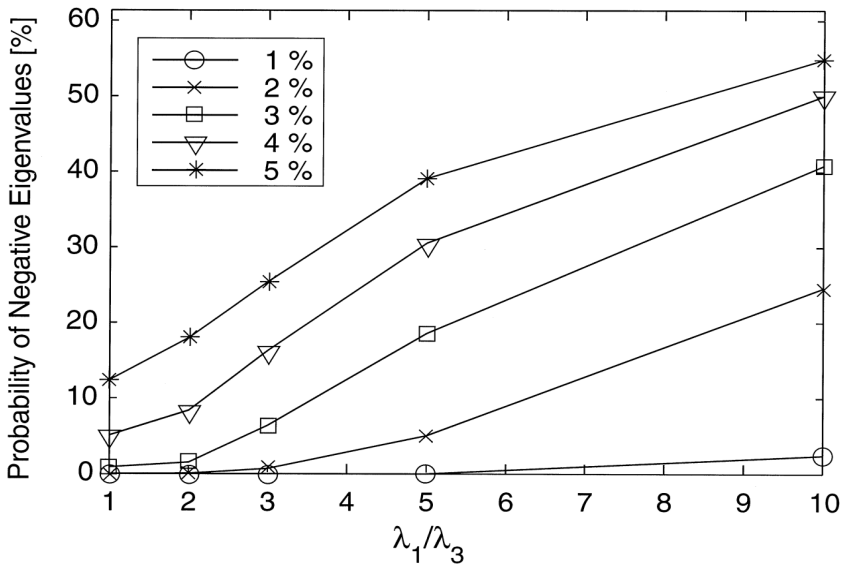


Figure 1.5: The probability of obtaining negative eigenvalues for the 'rice' shaped diffusion tensor as a function of λ_1/λ_3 at different noise levels (Skare et al., 2000).

λ_1/λ_3 for "rice" shaped diffusion tensors at different noise levels (%1 to %5) (Figure 1.5) [15]. Skare demonstrates that negative values happen more likely with increasing noise. Negative eigenvalues mainly occur in regions with high diffusion anisotropy such as the corpus callosum and the corticospinal tract due to more signal drop, and it may cause the FA, that is ranging from 0 to 1 theoretically, to exceed unity.

Several solutions are suggested in this regards. Some studies suggested to address noise in original DWI data. Using a moderate smoothing filter helps to increase signal to noise ratio (SNR) so that the risk of negative eigenvalues become negligible. Some other studies applied constrained estimation of diffusion tensors to ensure positive semidefinite (PSD). Koay et al. describes two linear and non-linear constrained least squares methods for eliminating the negative eigenvalues [16].

Neithammer et al. [17] restricts the estimated tensor to the PSD manifold using a gradient descent scheme for tensor estimation.

In this study, we enhanced `DTIProcess` and `DTIEstim` tools by applying the following suggestions to address negative eigenvalues issue:

- Correct eigenvalue measures after Tensor decomposition.
  - Calculate rotationally scalar values (RISs) with negative eigenvalues reset to zero.
  - Calculate RISs using absolute values of eigenvalues.

  A new flag named "–correction { zero/abs }" is added to `DTIProcess` and `DTIEstim` to avoid negative eigenvalues in DTI image before the construction of rotationally invariant scalars.
- Apply correction mechanisms to the estimated tensors to ensure positive semidefiniteness.
  - Correct the invalid estimated tensor $\mathbf{D}$ by finding its nearest symmetric positive semidefinite matrix [18].

  This method is incorporated to `DTIProcess` and `DTIEstim` tools by adding a new option called "–correction {nearest}".

Figure 1.6 compares the DTI estimation without any correction to the explained 3 correction methods added to `DTIEstim`. It looks like that "nearest" option is visually cleaner and seems more biologically plausible. Therefore, this option was used in our data processing pipeline [19].

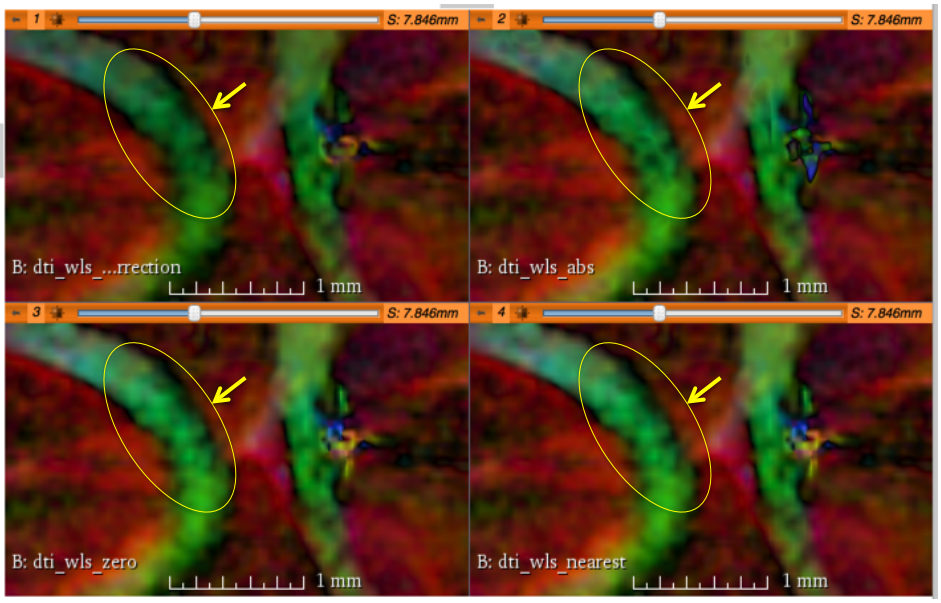


Figure 1.6: Comparison of tensor estimation with and without correction methods. Top left: results without correction. Top right: absolute value of eigenvalues are used. Bottom left: negative eigenvalues are reset to zero. Bottom right: nearest symmetric positive semidefinite matrix is found for each invalid tensor, which shows visually cleaner results.

Table 1.1: Summary of commonly used DTI scalars.

RIS Name	Formula	Interpretation
Fractional anisotropy	$FA = \sqrt{\frac{3[(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2]}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}}$	Shape of diffusion
Mean diffusivity	$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}$	Diffusion tensor trace average; size of diffusion
Radial diffusivity	$RD = \frac{\lambda_2 + \lambda_3}{2}$	Transverse diffusion
Axial diffusivity	$AD = \lambda_1$	Longitudinal diffusion

After correcting the invalid tensors, several rotationally invariant scalars (FA, MD, RD, and AD) were computed by `DTIProcess`. Based on the assumption of $\lambda_1 >= \lambda_2 >= \lambda_3 >= 0$, all RIS images now have all non-negative values.

1.4.3 Calculation of Rotationally Invariant Scalars (RISs)

Rotationally invariant scalars (RISs) are numeric representations of diffusion shape and magnitude [20] and are generated from resulting eigenvalue decomposition of diffusion tensors. They are important measures for clinicians and are useful biomarkers of disease progression. Four scalars derived from the tensor model of diffusion are often used and are referred as DTI scalars as well, including fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD), and axial diffusivity (AD). A summary of these commonly used DTI scalars is presented in Table (1.1) [14] .

FA reflects anisotropy of the diffusion ranging from 0 to 1. The MD (often called the apparent diffusion coefficient or ADC) is simply the trace of diffusion matrix divided by three, which is equivalent to the average of the eigenvalues [21]. FA and MD are commonly used measures of diffusivity because they summarize general shape and magnitude of diffusion, respectively, by accounting for diffusion magnitudes along three orthogonal directions at once [22]. FA is a summary measure of microstructural integrity. Larger axon diameter or lower packing density of axons both reduce the FA [23]. We expect to see lower FA in the areas where tract degeneration was expected to occur. However, reduced FA alone may not be able to fully capture changes in axonal integrity when changes in diffusion are subtle in one or two of the orthogonal directions. It may be more helpful to examine these subtle changes using directional measures of diffusivity to see diffusion magnitudes perpendicular and parallel to the first eigenvector [24]. MD is an inverse measure of the membrane density, so an increase of MD is expected where tract degeneration was expected to occur.

Axonal diffusivity (AD) is the magnitude of diffusion parallel to the principal direction of diffusion and decreases with axonal injury and degeneration [25]. Radial diffusivity (RD) is the magnitude of diffusion perpendicular to the principal direction of diffusion and increases with demyelination/dysmyelination [25, 26, 27]. When using measures of directional diffusivity (RD and AD), it is common to see changes in both RD and AD in a given region because the processes that cause changes in these measures (axonal death and demyelination) often occur in close proximity [25]. However, it is not always the case! An increase in RD with no change in AD has

been documented in a very specific type of myelin pathology called dysmyelination. Dysmyelination is the incomplete myelination of functional axons, as opposed to demyelination that is the loss of myelination [26, 27].

1.4.4 Units of Rotationally Invariant Scalars

Diffusion or Brownian motion refers to a random, microscopic movement of water and other small molecules due to their thermal energy at temperatures above zero. Based on Einstein diffusion equation [28] , the diffusion tensor $\mathbf{D}$ (equation 1.5) has the unit of (mm^2/s) proportional to the mean squared-displacement divided by the number of dimensions and the diffusion time.

One can think of that intuitively because diffusion is the flux of particles across a surface in a period of time, the units of $\mathbf{D}$ are of the form area/time.

The unit of the eigenvalue is the same as the unit of the matrix itself. This becomes evident from the eigenvalue equation:

$$\mathbf{D}v = \lambda v \tag{1.21}$$

Therefore, all the eigenvalues (λ_1,λ_2,λ_3) have the unit of (mm^2/s). As a result, mean diffusivity (MD), axial diffusivity (AD) and radial diffusivity (RD) have the unit of (mm^2/s), since they are computed from eigenvalues. However, Fractional Anisotropy (FA) is dimensionless ranging from 0 (isotropic diffusion) to 1 (high anisotropy) [20]. Also, the unit for Frobenius norm is the same as the unit of diffusion tensor matrix (mm^2/s) since Frobenius norm is directly computed from tensor matrix

elements:

$$\|\mathbf{D}\|_F = \sqrt{trace(\mathbf{D}^*\mathbf{D})} = \sqrt{\sum_{i=1}^{3}\sum_{j=1}^{3}|D_{ij}|^2} \tag{1.22}$$

Finally, note that all computed statistics (standard deviation, mean, maximum, minimum, and median) have the same unit as the data.

$e^{-b\mathbf{D}}$ will be dimensionless, the units of b-value should be the inverse of tensor $\mathbf{D}$. In other words, b should be expressed as [time/area]. Typical b-values available on modern MRI scanners range from 0 to about 4000 s/mm^2.

The optimal choice of b-value depends upon field strength, number of signals averaged, anatomical features, and predicted pathology. A higher b-value results in more diffusion weighting but also more noise. Mechanical vibration artifacts may also be a problem as b-values are increased. Although it is not clearly defined, most routine clinical DWI currently use b-values between 0 and 1000 [29].

1.4.5 Distance Metrics for Diffusion Tensors

One key factor in analysis of Diffusion Tensor Imaging (DTI) datasets is a proper choice of diffusion tensor distance that measures similarity or dissimilarity between the tensors. Several different measures have been proposed in the literature to compute distance between diffusion tensors. For two tensors $\mathbb{T}_1$ and $\mathbb{T}_2$, a distance measure refers to a function that has two tensors as inputs and returns a nonnegative scalar value:

$$d : Sym_3^+ \times Sym_3^+ \mapsto \mathbb{R}_0^+ \tag{1.23}$$

Where Sym_3^+ is a symmetric positive definite second-order tensor.

A distance measure d is a *metric* if, for the two tensors $\mathbb{T}_1$ and $\mathbb{T}_2$, it satisfies the following conditions [30]:

$$\mathbb{T}_1 = \mathbb{T}_2 \Leftrightarrow d(\mathbb{T}_1, \mathbb{T}_2) = 0 \tag{1.24}$$

$$d(\mathbb{T}_1, \mathbb{T}_2) = d(\mathbb{T}_2, \mathbb{T}_1) \tag{1.25}$$

A measure is a Riemannian metric, if it also fulfills the following condition for infinitesimally close $\mathbb{T}_1$ and $\mathbb{T}_2$:

$$d(\mathbb{T}_1, \mathbb{T}_2) \leq d(\mathbb{T}_1, \mathbb{T}_3) + d(\mathbb{T}_3, \mathbb{T}_2) \tag{1.26}$$

Diffusion tensors can be described by their different properties: size, orientation, and shape [30]. Distance measures based on scalar indices [31, 32] or angular differences [33] are invariant to one or more of the tensor properties; therefore, there are many tensors $\mathbb{T}_1 \neq \mathbb{T}_2$ for which $d(\mathbb{T}_1, \mathbb{T}_2) = 0$ that invalidates the metric condition (1.24). Scalar indices are computed using only rotationally invariant eigenvalues of the tensors, so they do not depict the directional variation of diffusion anisotropy. Also, angular measures only consider changes in orientation of the main eigenvalue, so they do not depict the size variation of diffusion tensors.

Here, we only consider the distance measures that are Riemannian metrics and fulfill the metric conditions (1.24) to (1.26). Following explains three well-known distance metrics.

1.4.5.1 L^2 Distance or Frobenius Distance

This measure deals with the diffusion tensors components as vector elements and computes the L^2-norm of the difference between the elements of two vectors.

$$d_{L2}(\mathbb{T}_1, \mathbb{T}_2) = \sqrt{\sum_{i=1}^{3}\sum_{j=1}^{3}(\mathbb{T}_{1ij} - \mathbb{T}_{2ij})^2} \tag{1.27}$$

d_{L2} is the same as the *Frobenius distance* [34] that is computed as:

$$d_F(\mathbb{T}_1, \mathbb{T}_2) = \sqrt{tr((\mathbb{T}_1 - \mathbb{T}_2)^2)} \tag{1.28}$$

where *tr(.)* is the matrix trace and is defined as the sum of the diagonal elements.

1.4.5.2 Riemannian Geometric-Based Distance

This class of measures constrains the matrices to positive definite matrices. Batchelor *et al.* [35] defined a geometric-based distance metric (d_g) between a pair of tensors and the associated shortest path (geodesic) joining them:

$$d_g(\mathbb{T}_1, \mathbb{T}_2) = N(\mathbb{T}_1^{-\frac{1}{2}}\mathbb{T}_2\mathbb{T}_1^{-\frac{1}{2}}) \tag{1.29}$$

where

$$N(\mathbb{T}) = \sqrt{\sum_{i=1}^{3}(log(\lambda_i^{\mathbb{T}}))^2} \tag{1.30}$$

where $\lambda_i^{\mathbb{T}}$ are the eigenvalues of tensor $\mathbb{T}$. This distance measure can also be used to compute tensor means or do interpolation between tensors [36, 37].

Additionally, this distance measure is invariant to any linear changes of coordinates and is also called "Affine-invariant" distance in some references [37, 38].

Affine-invariant Riemannian metrics have excellent theoretical properties, but also lead in practice to complex and slow algorithms. To remedy this limitation, Ar-

signy *et al.* [36] introduced a new family of Riemannian metrics called Log-Euclidian distance metric (d_{LE}). In their suggested framework, Riemannian computations can be converted into Euclidean ones once tensors have been transformed into their matrix logarithms. Therefore, d_{LE} is equivalent to the Frobenius distance of the logarithm of the matrices:

$$d_{LE}(\mathbb{T}_1, \mathbb{T}_2) = \sqrt{tr(\{log(\mathbb{T}_1) - log(\mathbb{T}_2)\}^2)} \tag{1.31}$$

where *tr(.)* is the matrix trace operator.

The Log-Euclidian distance yields same results as Riemannian distance, but with much simpler and faster computations [36]. Although Log-Euclidean metric does not yield full affine-invariance as the affine-invariant metric defined in equation (1.29), it is invariant by similarity (orthogonal transformation and scaling).

1.4.5.3 Kullback-Leibler Distance

A diffusion tensor can be interpreted as the covariance matrix of a Gaussian distribution that describes the local diffusion [30]. Therefore, a family of dissimilarity measures between diffusion tensors is the statistical divergence that measures the overlapping of probability distribution functions. In this regard, the square-root of the J-divergence (symmetrized Kullback-Leibler) is proposed as a distance metric (d_{KL}) between two diffusion tensors [39]:

$$d_{KL}(\mathbb{T}_1, \mathbb{T}_2) = \frac{1}{2}\sqrt{tr(\mathbb{T}_1^{-1}\mathbb{T}_2 + \mathbb{T}_2^{-1}\mathbb{T}_1) - 2n} \tag{1.32}$$

where the dimensionality n is 3 for diffusion tensors. Like Riemannian distance, d_{KL} is invariant to Affine transformations that is a desirable property in registration and

segmentation applications.

1.4.5.4 Comparison of Tensor Distance Metrics

Peeters *et al.* [30] have classified several distance and similarity measures that have been presented in literature. They have analyzed the behavior of above distance metrics according to their robustness to noise and their sensitivity to changes in size, orientation and shape of the diffusion tensors. The Shape is given by the ratio between the different eigenvalues, such that eigenvalues are changed from linear ($\lambda_1 > \lambda_2 = \lambda_3$) to planar shape ($\lambda_1 = \lambda_2 > \lambda_3$), and from planar to spherical shape ($\lambda_1 = \lambda_2 = \lambda_3$).

They have shown that all above-mentioned metrics are sensitive to changes in different properties of the diffusion tensor. All d_g, d_{LE} and d_{KL} give practically the same results and are sensitive to noise. The L^2 distance (d_{L2} or d_F) is relatively simple but shows higher robustness against the noise.

1.5 Partial Volume Effect (PVE)

Quantitative analysis of magnetic resonance (MR) brain images is important to advance our knowledge about human brain structures. In an MR acquisition, the different levels of energy emitted from different types of tissue are quantized and sampled into discrete pixel/voxel segments in an output image. However, due to the finite spatial resolution of the imaging device, a single image voxel may contain of several types of tissues. This phenomenon is called partial volume effect (PVE), and it complicates the segmentation process. Figure (1.7) shows a schematic explanation of the partial volume effect in the context of MR imaging. Voxels composed of purely

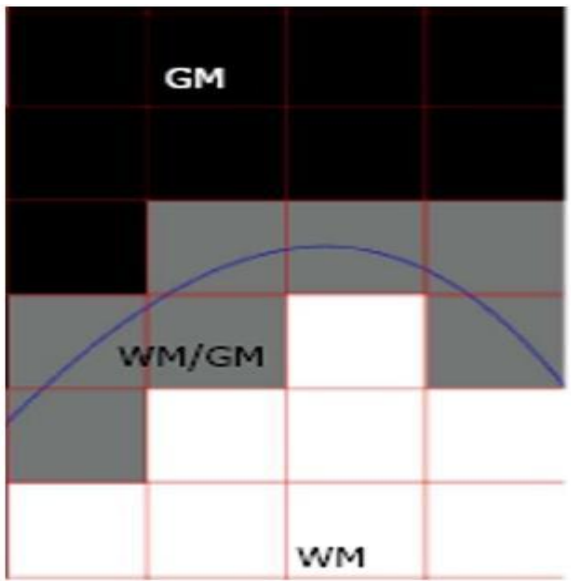


Figure 1.7: A schematic explanation of the partial volume effect in the context of brain magnetic resonance imaging (Tohka, 2014).

gray matter (GM) are colored in black while voxels composed of white matter (WM) are in white color. These are termed pure tissue voxels or pure plugs. Voxels composed of multiple tissue types, termed mixed voxels or non-pure plugs, are colored in gray. In the figure, these are voxels containing both WM and GM tissue types. The actual anatomical boundary between tissue types is shown in blue color, and red color is used to indicate voxel boundaries [40].

Dealing with PVE is an important factor for accurate classification of brain structures and volume quantification of MR brain images. A range of 20% to 60% of volume measurement errors is reported by ignoring the PVE [41, 42]. Several algorithms have tried to solve an extended version of the tissue classification problem to estimate the amount of each tissue type within each voxel. This extended problem has

been referred by several names as fuzzy segmentation, partial volume segmentation, tissue fraction estimation, and partial volume estimation. The extended problem computes the partial volume coefficients (PVCs) within each voxel, for example, a voxel contains 20% of WM, 80% of GM, and 0% of CSF. It should be noted that the partial volume coefficients are different from tissue type probabilities. Usually the Bayesian tissue classifiers produce more crisp tissue type maps than the partial volume estimation algorithms [40].

In this study, we suggested to enhance brain segmentation by using only pure voxels for initialization or training of a selected classification method. One way to find pure samples is to compute partial volume coefficients (PVCs) within each voxel and choose only those voxels that demonstrate one tissue type is significantly dominant. *Mixed model* is the most commonly used model of partial volume estimation in brain MRI [43]. This model assumes that the intensity value of each voxel is a realization of a weighted sum of random variables (RVs), that each RV characterizes a pure tissue type. The weights are partial volume coefficients (PVCs) in the range of 0 to 1.

Consider the observed image as $X = \{\chi_i : i = 1, ..., N\}$, with $\chi_i \in R^K$, where N denotes the number of voxels in the image, i is the voxel index, and K is the number of modality channels in the multispectral case. A voxel intensity χ_i is considered as the realization of random variable χ_i. Assume there are M tissue types in the image; similarly, each tissue type j is described by a random variable l_j, that is assumed to be distributed according to a multivariate normal distribution with the mean of μ_j and covariance of Σ_j. Then, random variable χ_i can be written as the weighted sum

of tissue types:

$$\chi_i = \sum_{j=1}^{M} w_{ij} l_j + n \tag{1.33}$$

Where n represents the measurement noise that is assumed to be Gaussian, and $w_{ij} \in [0, 1]$ are partial volume coefficients (PVCs) for all tissues j in voxel i, such that in each voxel: $\sum_{j=1}^{M} w_{ij} = 1$.

The mixed model needs to be simplified. One simplification method is sampling noise model [44, 45], that assumes all randomness in the model is due to measurement noise. This leads to a new model where tissue types are represented by the mean intensities of tissue types:

$$\chi_i = \sum_{j=1}^{M} w_{ij} \mu_j + n \tag{1.34}$$

If enough data channels are available, a direct solution to above equation is through penalized least squares [43] my minimizing:

$$LS(W) = \sum_{i=1}^{N} \left\| \chi_i - \sum_{j=1}^{M} w_{ij} \mu_j \right\|^2 \tag{1.35}$$

with constrains that $\sum_{j=1}^{M} w_{ij} = 1$ and $0 \leqslant w_{ij} \leqslant 1$, and W denotes a matrix of all PVCs. A variation to above minimization problem is suggested by adding a regularization term with a Markov Random Fields prior that assumes PVCs of neighboring voxels should have similar values [43].

An accurate solution to partial volume estimation is possible if enough multi-channel information are available from different modality sources collected in high enough resolution. The accuracy of mixed model decreases if multi-modal information are coming from different spatial resolutions, where some modality images are

acquired in lower resolution than others. This makes the estimation of partial volume coefficients inaccurate especially in tissue boundaries. In chapter 3, Figure (3.5), we illustrated that incorporating multi-modal information, collected in different spatial resolutions, increases the number of mixed samples in tissue boundaries.

This work, in chapter 3, emphasizes that careful considerations should be taken to deal with increased PVE when input modality data are provided in different spatial resolutions. We introduced a novel method to find pure samples. Our method operates in physical spatial domain and is not limited by the constrains of voxel lattice spaces of different input modalities. Once detecting pure samples, we can safely integrate multi-modal information in training/initialization of selected classifiers for an enhanced segmentation quality.

PVE are consequences of limited spatial resolution and are increased in low-resolution images like diffusion-weighted MRI. In chapter 4, we addressed PVE in low-resolution DWI scans by introducing a novel super-resolution reconstruction approach using the multi-modal information of structural MR images provided in high spatial resolution. There are some other multi-modal PVE corrections introduced in literature. Shidahara *et al.* [46] introduced a synergistic functional-structural resolution recovery (SFS-RR) method for PVE correction at PET-MR. They proposed the use of a probability brain atlas by segmenting T1-MR, and run an ROI wavelet-based resolution recovery. Grecchi *et al.* [47] has developed the SFS-RR algorithm to fit PET-CT for detecting and monitoring bone metastases. Boussion *et al.* [48] introduced another multi-resolution approach to extract information from high-resolution

(HR) image (MR or CT) to transfer and integrate them in low-resolution image (PET or SPECT) using a discrete wavelet transform. In this study, for the first time, we suggested to use HR structural MR modalities for super-resolution reconstruction (SRR) of DWI scans by developing a weighted total variation (WTV) method to fit our super-resolution approach.

1.6 PVE Challenges in Processing of Large-scale Heterogeneous Data

The overall purpose of this research is to address limitations caused by partial volume effects (PVE) in processing of large-scale heterogeneous data. Dealing with PVE in an appropriate manner is important to get better interpretation from thousands of heterogeneous data scans acquired over the years for multi-site studies like the Huntingtons Disease study in the PREDICT-HD project [1].

There are hundreds of structural MR data, diffusion-weighted imaging data and clinical data from hundreds of controls and subjects with Huntingtin gene expansion acquired from several collection sites all over the world. It is desirable to use information from all acquired heterogeneous data, but it is not possible to naïvely incorporate them together in available standard processing tools without addressing current limitations. Over the first 7 years of PREDICT-HD study, most of the data were collected from scanners with 1.5 Tesla scan protocol, where the T2-weighted scans were acquired in lower spatial resolution than the T1-weighted scans typically by a factor of 3. This 7 years of low-resolution T2-weighted scans were ignored in our current processing pipeline as they were adversely affecting the quality of seg-

mentation results when they were directly processed by our standard segmentation algorithm. Also, hundreds of diffusion-weighted imaging (DWI) data were acquired in very low spatial resolution of 8 mm^3 that could provide complementary information in study of WM abnormalities and could be used to develop new biomarkers for disease progression [19]. However, application of DWI is limited by its relatively low spatial resolution.

In this study, we investigated the issues caused by partial volume effects due to the low-resolution nature of data, and we introduced novel techniques to tackle current limitations to combine complementary information from all low-resolution and high-resolution multi-modal scans acquired in a data session to advanced our interpretation of underlying anatomical structures.

CHAPTER 2
TISSUE CLASSIFICATION OF LARGE-SCALE MULTI-SITE HETEROGENEOUS MR DATA USING FUZZY K-NEAREST NEIGHBOR METHOD

This study describes enhancements to automate classification of brain tissues for multi-site degenerative magnetic resonance imaging (MRI) data analysis. Processing of large collections of MR images is a key research technique to advance our understanding of human brain. Previous studies have developed a robust multi-modal tool for automated tissue classification of large-scale data based on expectation maximization (EM) method initialized by group-wise prior probability distributions. This study aims to augment the EM-based classification using a non-parametric fuzzy k-Nearest Neighbor (k-NN) classifier that can model the unique anatomical states of each subject in the study of degenerative diseases. The developed method is applicable to multi-center heterogeneous data analysis.

2.1 Introduction

Brain tissue segmentation on structural magnetic resonance imaging (MRI) has received considerable attention; one of classic neuroimaging challenges is the segmentation of MR images into white matter (WM), grey matter (GM) and cerebrospinal fluid (CSF). Volumetric measurements in different brain regions are important in studies on aging and neurodegenerative disorders [49] like Alzheimer's disease, Schizophrenia and Huntington's Disease (HD).

Given the relevance of brain tissue segmentation, different automated segmen-

tation methods have been proposed over the years. Almost all of these methods rely on a supervised or unsupervised voxel classifier. Supervised methods use manually segmented training data to learn the typical distribution of intensity or appearance features for the tissue classes [50]. Unsupervised methods, particularly those based on expectation maximization (EM), do not require training data and are therefore more widely used than the supervised methods. EM-based methods start with an initial segmentation, that is often based on a probabilistic brain tissue atlas that is registered to the unlabeled target scans, and from this initialization, class-specific Gaussian intensity distributions are estimated. This intensity model can then be used to update the segmentation and this process is repeated until the segmentation converges [49].

Kim and Johnson [8] implemented an iterative optimization framework between bias-correction, registration, and tissue classification using expectation maximization (EM) method for large-scale heterogeneous multi-site longitudinal MR data analysis. In this study, we extend the EM-based classification using a non-parametric fuzzy k-Nearest Neighbor (k-NN) classifier that avoids biases inherent in EM use of prior probability distributions that may not represent diseased anatomical states.

2.2 Methods

2.2.1 General Framework

This study describes the algorithmic enhancements on the implementation of a framework developed by Kim and Johnson [8] that iteratively incorporates bias-

field correction, image registration, and tissue classification. Enhancements applied for more accurate subject specific tissue classification in processing of heterogeneous multi-site degenerative MR data.

Our atlas based framework takes inputs of any combination of modalities with any number of scan repetitions if the input modalities have comparable resolution and voxel sizes. First, using a Rigid-type registration, all intra-subject scans are spatially normalized into a common subject-specific reference orientation [51] defined by anterior commissure (AC), and posterior commissure (PC) landmarks, and mid-saggital plane [52, 53]. Then, all intra-modal scan repetitions are averaged together to increase the signal-to-noise ratio for each modality. After that, all the atlas priors are placed into the subject space using an atlas to subject transformation that is derived from a high-deformable registration algorithm (SyN) [54, 55] to enhance the accuracy of subject-specific tissue priors.

The warped priors in the subject space are tissue probability maps giving the probability of a certain voxel belonging to a certain tissue, and they are used to initialize the Gaussian distribution parameters for the EM algorithm. Finally, the processes of posterior estimation, bias field correction, and the registration are iteratively updated multiple times until convergence (Figure 2.1).

The core implementation of the framework developed by Kim and Johnson [8] uses a general expectation-maximization (EM) algorithm [56, 57, 58, 59] by iterative distributional parameter estimation to classify each individual voxel y at location i into K tissue types. The process assumes Gaussian mixture model $y_i \sim N(\theta_i)$ where

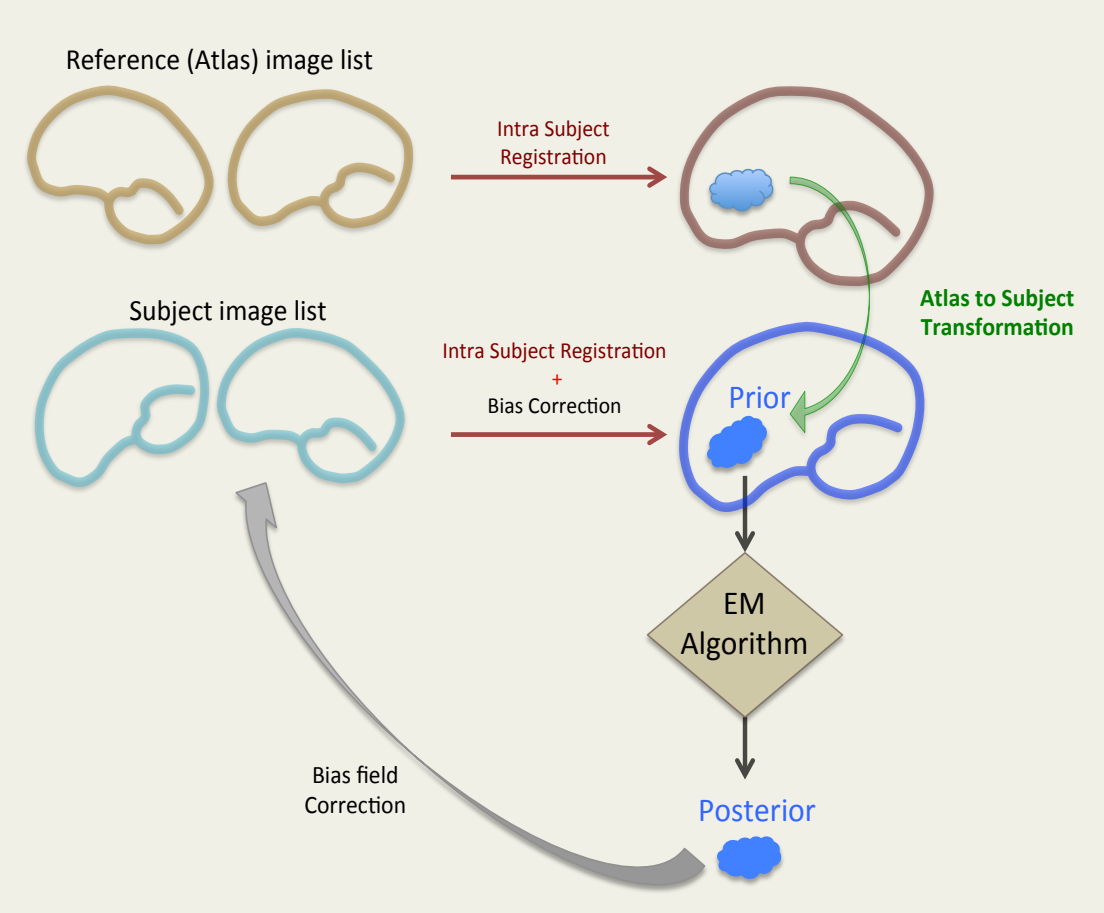


Figure 2.1: A rough demonstration of iterative classification process by (Kim and Johnson, 2013) using expectation maximization (EM) method. In our development, the EM output posterior probability maps and the input multi-modal MR images are used to train a fuzzy k-NN classifier.

$\theta_i = \{\mu_i, \sigma_i\}$ with mean μ and variance σ^2 of each tissue label $\Gamma \in \{l | l = 1 \cdots K\}$. First step is the expectation (E) step to determine the expected posterior density function $p(y_i|\theta, \Phi_i)$ that tells us how likely voxel y_i belongs to each tissue label with estimated bias-field Φ_i [57]. Formula presented here are mostly taken from the works [8, 56, 57, 58, 59].

E-Step.

$$p(y_i|\theta, \Phi_i) = \sum_l p(y_i|\Gamma_i = l, \theta_l, \Phi_i) p(\Gamma_i = l), \tag{2.1}$$

with $p(y_i|\Gamma_i = l, \theta_l, \Phi_i) = N_{\sigma,l}(y_i - \mu_l - \Phi_i)$ and

$$p(\Gamma_i = l) = \frac{t_{il}}{\sum_{l=1}^{k} t_{il}}, \text{where } t_{il} \text{ is tissue specific prior} \tag{2.2}$$

Then the maximization (M) step updates the parameters of Gaussian θ and bias-field Φ by maximum likelihood estimation from current density function.

M-Step.

$$\mu_l = \frac{\sum_i p(\Gamma_i = l|y_i, \theta, \Phi_i)(y_i - \Phi_i)}{\sum_i p(\Gamma_i = l|y_i, \theta, \Phi_i)}, \sigma_l^2 = \frac{\sum_i p(\Gamma_i = l|y_i, \theta, \Phi_i)(y_i - \mu_l - \Phi_i)^2}{\sum_i p(\Gamma_i = l|y_i, \theta, \Phi_i)} \tag{2.3}$$

Above equations are extended to multi-modal data as described in [57].

2.2.2 New Classifier

A non-parametric subject specific fuzzy $k-NN$ classifier complements the EM estimate of tissues using the information from multi-modal scans. The new classifier takes the output tissue probability maps (TPMs), $P_c(i)$, from the EM algorithm,

where "i" represents a voxel location, and "c" is a single tissue class:

$$\forall i \in \{voxel \quad locations\},$$

$$\forall c \in \{1, \ldots, \mathbb{C}\}, \tag{2.4}$$

$$\exists \quad 0 \leq P_c(i) \leq 1 \quad s.t. \quad \sum_{c=1}^{\mathbb{C}} P_c(i) = 1$$

Where $\mathbb{C}$ is the total number of tissue types, and $P_c(i)$ represent how likely voxel location i belongs to tissue type c.

2.2.2.1 Training Sample Set

To find the candidate training sample locations "t" for the $k-NN$ classifier, all $P_c(i)$ from EM posterior TPMs are thresholded in order to identify those sample locations that have a sufficient probability to belong to a single tissue type. Increasing the threshold leads to fewer but more reliable tissue samples. A threshold of 0.7 is chosen based on the results presented by Vrooman *et al.*[60] for brains tissue types.

$$t \in \{i \quad | \quad \exists c \quad s.t. \quad P_c(i) \geq 0.7\} \tag{2.5}$$

Where training sample location t is assigned with a label pointing to tissue region c.

Chosen training samples are then represented in an $\mathbb{F}$-dimensional feature space with:

$$\mathbb{F} = \mathbb{M} + \mathbb{C} \tag{2.6}$$

Where $\mathbb{F}$ is the number of features; $\mathbb{M}$ is the number of input multi-modal scans, and $\mathbb{C}$ is the total number of tissue types. The feature vector corresponding to the training sample t is created as:

$$\left[I_1(t), \quad ..., \quad I_{\mathbb{M}}(t), \quad \breve{P}_1(t), \quad ..., \quad \breve{P}_{\mathbb{C}}(t)\right] \tag{2.7}$$

Where $I_m(t)$, $m \in \{1, \ldots, \mathbb{M}\}$ represents the intensity value of m^{th} input image scan at sample location t, and $\breve{P}_c(t)$, $c \in \{1, \ldots, \mathbb{C}\}$ is a binary value derived from the c^{th} EM posterior TPM at sample location t, such that:

$$\breve{P}_c(t) = \begin{cases} 1 & if \quad P_c(t) \geq 0.01 \\ 0 & if \quad P_c(t) < 0.01 \end{cases} \tag{2.8}$$

In fact, our feature space defines all the candidate regions, suggested by EM results, that the current sample location t probably belongs to by more than one percent chance. In this way, the fuzzy $K - NN$ classifier is restricted to only biological plausible results, and it is not biased by the probability values estimated in EM step.

Finally, each created feature vector is added to a *training sample set*, and its known label code is added to the corresponding row of a *labels vector*.

2.2.2.2 Test Sample Set

Test sample locations "s" are the center points of voxel locations in the first input scan, and for each test location, a feature vector is created as shown in equation (2.7). All the test feature vectors are then added to a *test sample set*.

2.2.2.3 Run the Algorithm

The training and test sample sets and the labels vector are passed to a fuzzy $k - NN$ algorithm where the following procedure is performed on each test sample location:

1. In the feature space, the Euclidean distances between each test sample and all the training samples are computed. Distances are calculated through a k-

dimensional tree structure [61] that is a data structure for organizing points in a k-dimensional space using space partitioning.

2. The first $\mathbb{K}$ nearest neighbors are identified from the computed distance vector. $\mathbb{K}$ needs to be an odd number, and it was set to 45 as suggested by Vrooman *et al.* [60] and Cocosco *et al.* [62].

3. New probabilities, $P_c(s)$, are computed for the test location s showing how likely the current test location belongs to each tissue type. If $\mathbb{N}$ out of $\mathbb{K}$ nearest neighbors belong to tissue class c, then:

$$\begin{gathered} \forall s \in \{test \quad sample \quad locations\}, \\ \forall c \in \{1, \ldots, \mathbb{C}\}, \\ P_c(s) = \frac{\sum_{o=1}^{\mathbb{N}} \frac{1}{d_{c,o}^2}}{\sum_{i=1}^{\mathbb{K}} \frac{1}{d_i^2}} \end{gathered} \tag{2.9}$$

Where $d_{c,o}$ is the distance of o^{th} occurrence of class c to the current test location s; d_i is the distance to the i^{th} neighbor of current test sample, and $P_c(s)$ represents the probability that the current test location s belongs to class c.

4. For the test location s, all computed $P_c(s)$, $c \in \{1, \ldots, \mathbb{C}\}$ are stored in one **row** of a $\mathbb{S} \times \mathbb{C}$ *likelihood* matrix, where $\mathbb{S}$ is the total number of test locations, and $\mathbb{C}$ is the total number of tissue types:

$$\text{likelihood matrix} = \begin{bmatrix} & & \vdots & & \\ P_1(s), & \ldots, & P_c(s), & \ldots, & P_{\mathbb{C}}(s) \\ & & \vdots & & \end{bmatrix}_{\mathbb{S} \times \mathbb{C}} \tag{2.10}$$

5. Finally, new tissue probability maps are created by rearranging each **column**

of likelihood matrix to an output probability image. There are $\mathbb{C}$ output probability maps corresponding to all interested tissue types.

2.3 Experimental Methods

The accuracy and effectiveness of developed method is evaluated qualitatively and quantitatively.

2.3.1 Test Data

A set of 18 synthetic MR datasets of a brain subject from BrainWeb database [63] are used for quantitative evaluation of developed enhancements. The BrainWeb database provides a rich set of multi-spectral data as *input sources* to our algorithm that include both $T1$ and $T2$ modality scans. BrainWeb also provides a simulation of heterogeneous nature of multi-site real data with input variants that represent six levels of noise and three degrees of bias-field for each $T1$ and $T2$. Finally, the BrainWeb data provides a set of tissue segmentation *baselines* for comparison against each *output result* from our algorithm.

2.3.2 Qualitative Evaluation

3D Slicer [64] was used to visually compare the segmentation results of developed enhancements to the technique established by Kim and Johnson [8]. Qualitative investigation was done using a sample $T1$-weighted MR scan that was arbitrarily selected from our local University of Iowa SIEMENS Trio Tim 3 Tesla scan protocol. This protocol was used as part of multi-site international PREDICT-HD [1] project.

Table 2.1: Atlas definition of 15 region-specific intensity-context priors. Each tissue type is sub-divided into regions of interest with given names. (Gm = Gray matter, Wm = White matter, Csf = Cerebrospinal fluid, Crbl = Cerebellum, Vb = Venous blood).

Tissue	Gm	Wm	Csf	Wm & Gm	Vb	Background
Name	Basal Hippocampus Crbl Gm Surf Gm	Wm Crbl Wm	Csf	Thalamus Globus	Vb	Not Gm Not Wm Not Csf Not Vb Air

2.3.3 Quantitative Evaluation

Developed enhancements were evaluated quantitatively using BrainWeb database described in section 2.3.1.

The accuracy and robustness of developed enhancements by a fuzzy $k-NN$ algorithm were compared to the reported results by Kim and Johnson [8] derived from an EM-based only classification. For this purpose, the similarity of both methods were compared against the segmentation baseline provided by BrainWeb along with the evaluation datasets.

Our atlas based approach uses the atlas definitions from two T1-weighted and T2-weighted modalities with priors for 15 discrete region-specific tissue types listed in Table 2.1. This is a slightly modified approach to that taken in [8] where 17 regions were identified with the Basal region being subdivided into (Caudate, Putamen, Accumben) regions.

The software is implemented based on the Insight Toolkit (ITK) libraries [65, 66] and conforms to the coding style, testing, and software guidelines identi-

fied by the National Alliance for Medical Image Computing (NAMIC) group. Our implementation is publicly available via BRAINSTools package [67] and contributes to a fully automated processing pipeline for MR images [68, 69].

2.4 Results

Figure 2.2 shows the visual comparison of results on a sample MR scan from the PREDICT-HD study. As shown by corresponding arrows in both images, the segmentation boundaries of GM, WM and CSF from our developed method (using k-NN) (Fig. 2.2(b)) are more agreeable to real anatomical tissue boundaries than the results derived from EM-based only classification (Fig. 2.2(a)).

In order to compare the quantitative results, two independent measures, "Dice index" and "average Hausdorff distance [70]", are reported to compare the results of automated delineations against the ground truth. Figure 2.3 shows the Dice index (larger is better) and average Hausdorff distance (smaller is better) evaluated along three degrees of bias-field ($rf = 0\%$, $rf = 20\%$ and $rf = 40\%$) and six levels of noise (0%, 1%, 3%, 5%, 7%, and 9%) for three tissue types (WM, GM and CSF). The results of EM method are shown in *black* while the *blue* color is used for the results of developed enhancements using a k-NN classification [71].

2.5 Discussion and Conclusions

This research project improved automated classification of brain tissues for multi-center $3D$ MRI data analysis. Previous studies have used expectation maximization (EM) based classification that is group specific and uses a *priori* knowledge

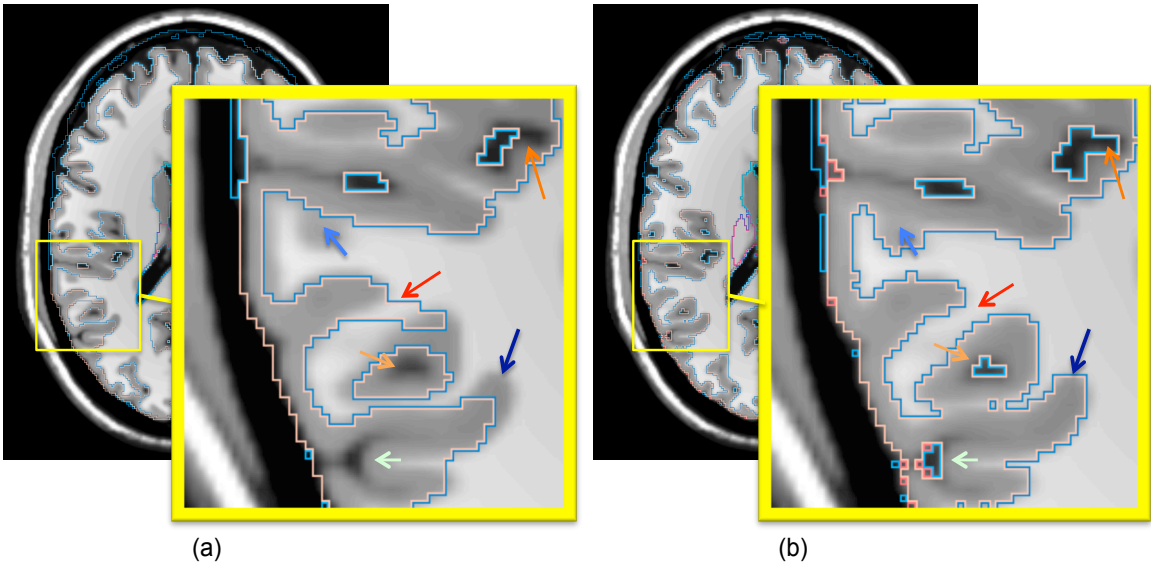


Figure 2.2: Visual comparison of segmentation results on a sample PREDICT-HD MR data between (a) EM only based classification and (b) developed enhancements using a fuzzy $k-NN$ classifier. More accurate delineation is achieved by the developed method.

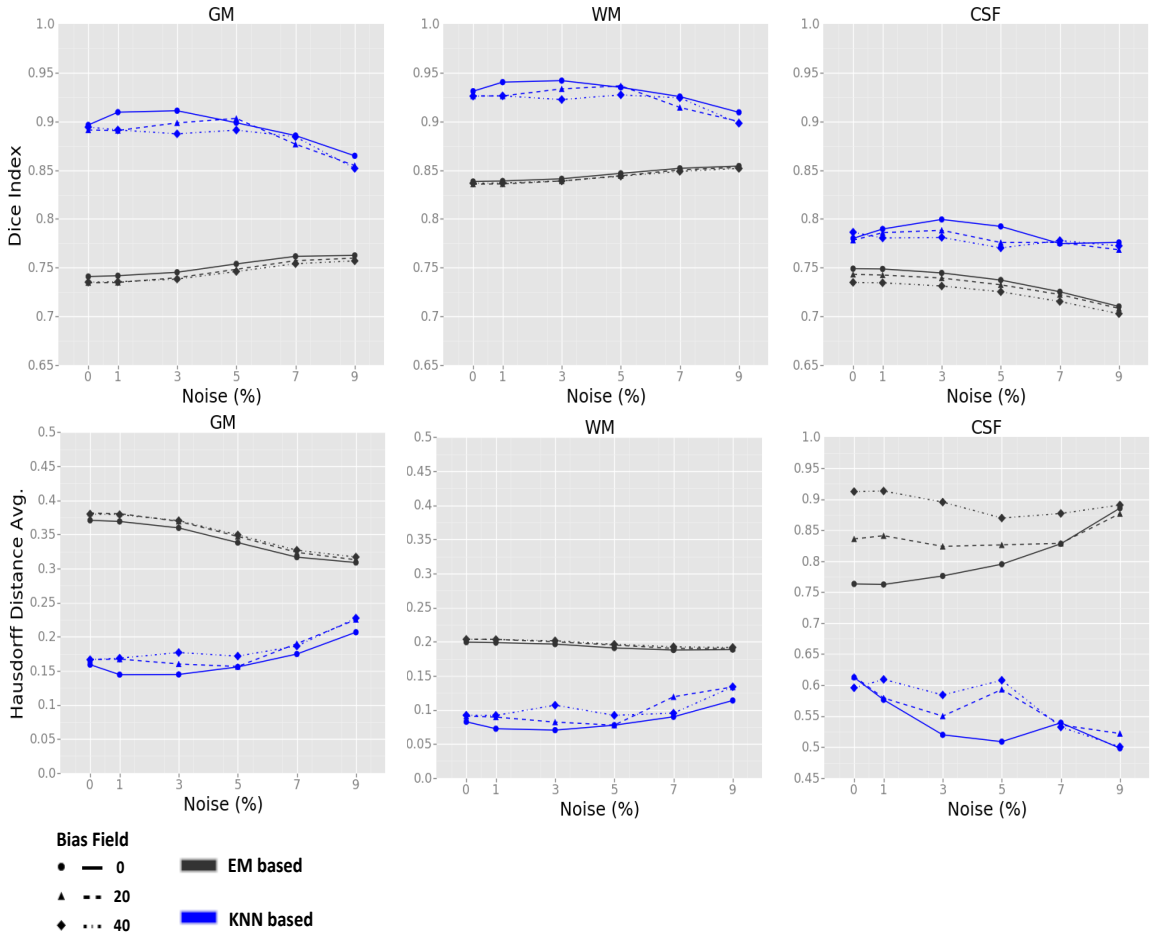


Figure 2.3: Comparison of classification of cerebrospinal fluid (CSF), Grey matter (GM) and White matter (WM) tissues between EM only based classification (black) and the extended method by a fuzzy $k-NN$ classifier (blue) using two independent measures, Dice index (larger is better) and average Hausdorff distance (smaller is better) . The evaluation is performed along three degrees of bias-field (rf=0, rf=20 and rf=40) and six levels of noise (0%, 1%, 3%, 5%, 7%, and 9%) along x-axis. Both similarity measures show improvement on the results of developed $k-NN$ enhancements.

for all the subjects in an atlas based approach. This study, however, emphasized the importance of a non-parametric model's utility in neurodegenerative diseases, since each subject has unique anatomical states in longitudinal degenerative studies that may not be represented by prior probability distributions. Enhancements were suggested by augmenting the EM-based classification using a fuzzy k-nearest neighbor (k-NN) classifier that builds up a model for each individual subject and complements the classification results that EM produces. A Fuzzy $k-NN$ method was selected, as this non-parametric classifier is subject specific; it is not biased by prior probability distributions, and it potentially can model more complex decision boundaries than a Gaussian distribution based mixture method.

The developed method generates more precise results than EM only classification since both similarity measures, Dice index and the average Hausdorff distance, show improvement on the results of $k-NN$ classifier as demonstrated in Figure 2.3. Also, qualitative observations in Figure 2.2 show that our method especially provides more accuracy in delineation of sophisticated tissue boundaries where tissue regions are highly interleaved together like GM and WM boundaries.

Also, to investigate more about the feature vector of the $k-NN$ model (equation 2.7), we compared the results of $k-NN$ classification against the ground truth for each of following approaches:

- Feature vectors were created using the intensity values of the input images and the values of EM tissue probability maps.
- Feature vectors were created using the intensity values of the input images and

thresholded values of EM tissue probability maps (as described in equation 2.8).

Figure (2.4) compares the above approaches using two independent metrics, Dice index (larger is better) and average Hausdorff distance (smaller is better), evaluated along three degrees of bias-field ($rf = 0\%$, $rf = 20\%$ and $rf = 40\%$) and *six* levels of noise (0%, 1%, 3%, 5%, 7%, and 9%) for three tissue types (WM, GM and CSF). The results of described training method using thresholded EM tissue probability maps (equation 2.8) are shown in *blue* while the *red* color shows the results when $k-NN$ was trained based on direct values of EM tissue probability maps. The training approach using thresholded values (as in equation 2.8) generates more precise results because the training of $k-NN$ method is not biased by the probability values estimated in EM step, but the $k-NN$ model still benefits from the EM results as it is restricted to only candidate regions suggested by the EM outputs.

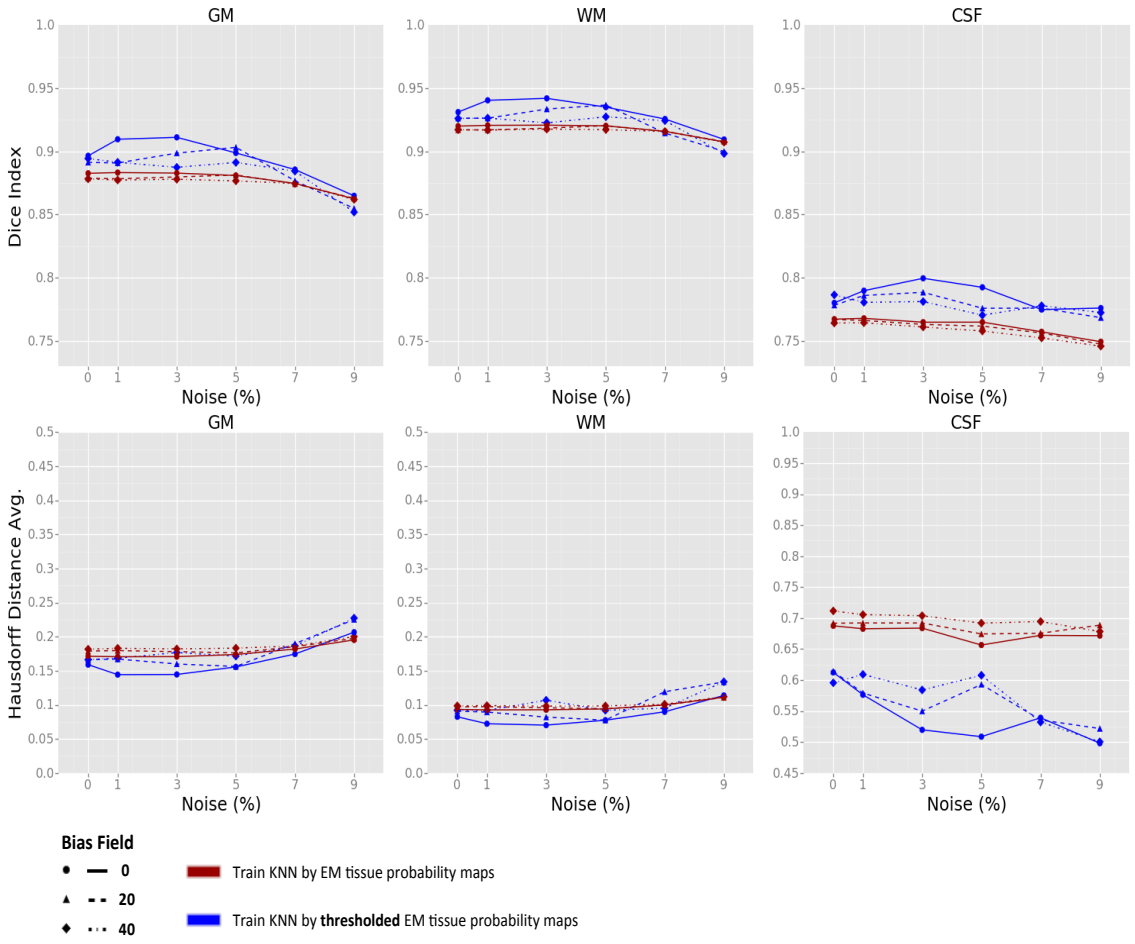


Figure 2.4: Two independent measures, Dice index (larger is better) and average Hausdorff distance (smaller is better), are used to compare the classification of cerebrospinal fluid (CSF), Grey matter (GM) and White matter (WM) tissues in two cases: (1) $k-NN$ classifier was trained by thresholded EM tissue probability maps (equation 2.8) (results in blue); (2) $k-NN$ classifier was trained directly by EM tissue probability maps (results in red). The evaluation is performed along three degrees of bias-field (rf=0, rf=20 and rf=40) and six levels of noise (0%, 1%, 3%, 5%, 7%, and 9%) along x-axis. Both similarity measures show improvement on using the thresholded EM tissue probability maps in the training phase of $k-NN$.

CHAPTER 3
ENHANCE MR MULTI-MODAL TISSUE CLASSIFICATION BY ADDRESSING PARTIAL VOLUME EFFECT

Previous chapter suggested enhancements when multi-modal MR modalities ($T1/T2$-weighted) were acquired at the *same* high spatial resolution with isotropic 1 mm^3 voxel sizes. However, many real world data are collected such that $T2$ image is acquired at lower spatial resolution than $T1$ (usually by a factor of 2 to 3). This is especially the case in datasets provided from scanners with 1.5 Tesla scan protocol. This chapter aims to enhance multi-modal classification when complementary information comes from multiple modality scans with different spatial resolutions.

At the first step, it is important to upgrade the previous multi-modal classification framework, introduced in chapter 2, to run in physical space before we investigate the segmentation results in the case that input modality scans are not at the same resolution and their voxel lattices do not line up.

In this chapter, first we enhance our previous multi-modal classification framework to run in physical space, and we demonstrate that the system is upgraded successfully by generating equivalent results in voxel space and physical space in the case of single-modal input. Then, we show that naïvely adding the information of a second modality with lower spatial resolution can adversely affect the segmentation results. We investigate and explain the reason by describing partial volume effects (PVE), and we develop a novel approach to deal with PVE issue. Finally, we present the evaluation results of developed method.

3.1 Enhancement of Classification Framework to Run in Physical Space

Different modalities may have different spatial resolutions in a multi-modal classification framework. Therefore, it is important to perform classification in physical space to avoid interpolation errors that may introduce artificial partial volume effects. In this step, the previous classification framework, introduced in chapter 2, is reimplemented to perform tissue classification in physical space.

3.1.1 Evaluation

Only one modality (T1-weighted) is used to evaluate the performance of reimplemented framework. By using only one modality, we expect to get essentially the same results when the classification is run in physical space versus voxel space.

3.1.1.1 Test Data

BrainWeb dataset, as described in section 2.3.1, is used for quantitative evaluation of the classification performance.

3.1.1.2 Results

"Dice Index" measure is reported in Figure (3.1) to compare the results of automated delineations against the ground truth when the algorithm is run in physical space vs. voxel space.

3.1.2 Conclusion

Single-modal results show that there is no significant difference in the performance of the classification framework when the algorithm was run in physical space

rather than voxel space. That is our desirable, since it means the results are not deteriorated when the framework was upgraded. Now we are ready to do further investigations on multi-modal classification using the new proper infrastructure.

3.2 Significance of Multi-modal Classification

Using complementary information from different modality sources helps better delineation of different brain tissue types, since each modality scan reveals different characteristics of underlying biological structure. T1-weighted MR is the most used modality for brain segmentation; however, as an example, Figure (3.2 a) shows that it is hard to distinguish CSF from the bone marrow, or the gray matter (GM) from the venous blood (VB) by looking only at T1 image. As illustrated by Figure (3.2 b), a T2-weighted scan provides better contrast to resolve the ambiguities between above-mentioned biological tissue types.

Intracranial volume (ICV) is used as a statistical correction of overall body size in many neuroimaging structural analysis [72, 73]. Accurate segmentation of marginal brain regions like CSF and Venous Blood results in more accurate estimation of total ICV.

Figure (3.3) qualitatively shows that an accurate segmentation of CSF can be achieved by using complementary information from different modalities. We used a sample MR dataset that was arbitrarily selected from our local University of Iowa SIEMENS Trio Tim 3 Tesla scan protocol. This protocol was used as part of multi-site international PREDICT-HD project.

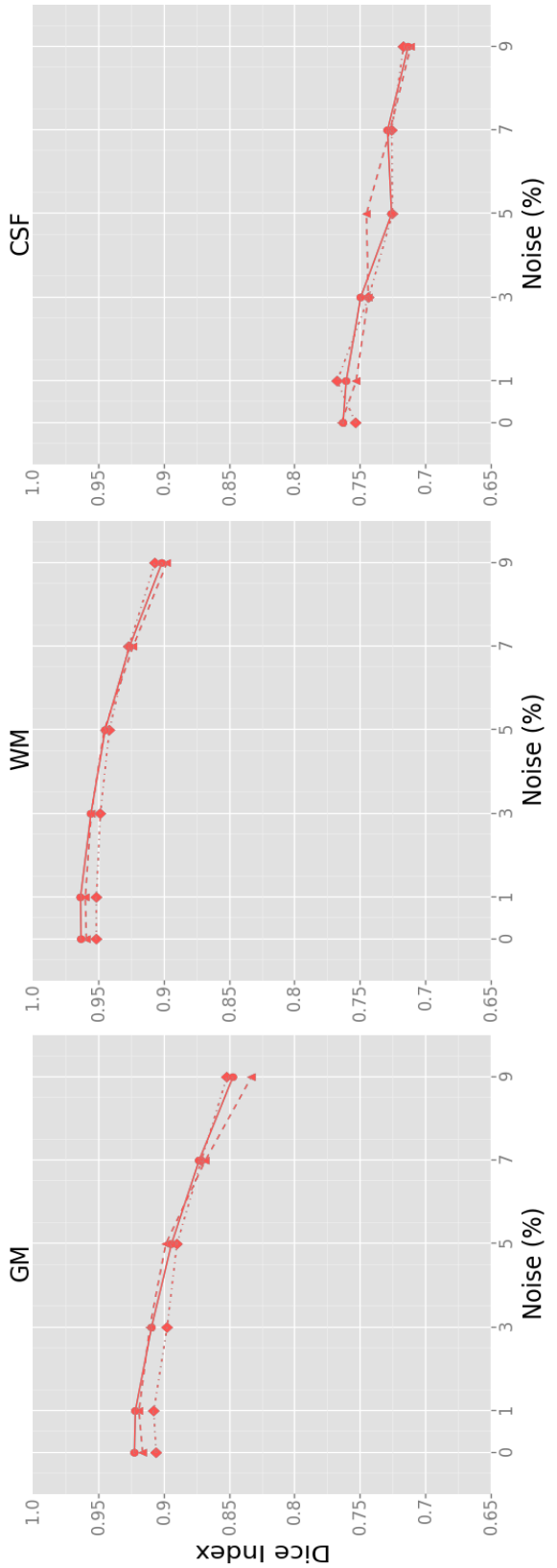


Figure 3.1: The classification framework generates essentially the same results in single-modal mode when it is run in physical space versus voxel space. Blue lines are masked by the red lines as the results are overlapped.

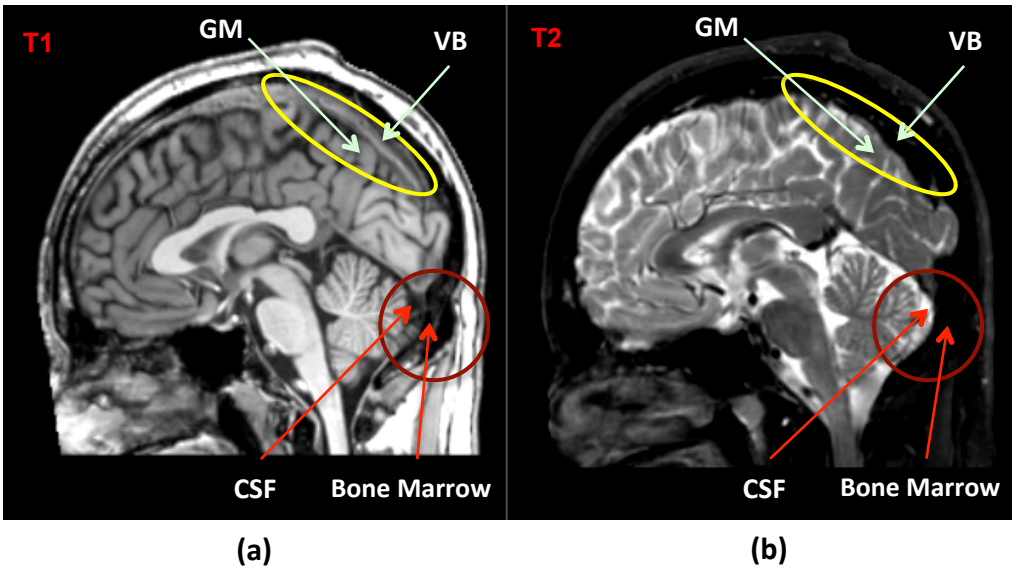


Figure 3.2: (a) It is hard to distinguish between gray matter (GM) and venous blood (VB) (as both are shown as grey) and to distinguish CSF from the bone marrow (since both are shown as dark) by looking only at T1-weighted MR scan. (b) T2-weighted MR provides better contrast in the case of T1 ambiguities between mentioned tissue types, since here GM is gray but VB is dark; also CSF is shown as white while bone marrow is dark.

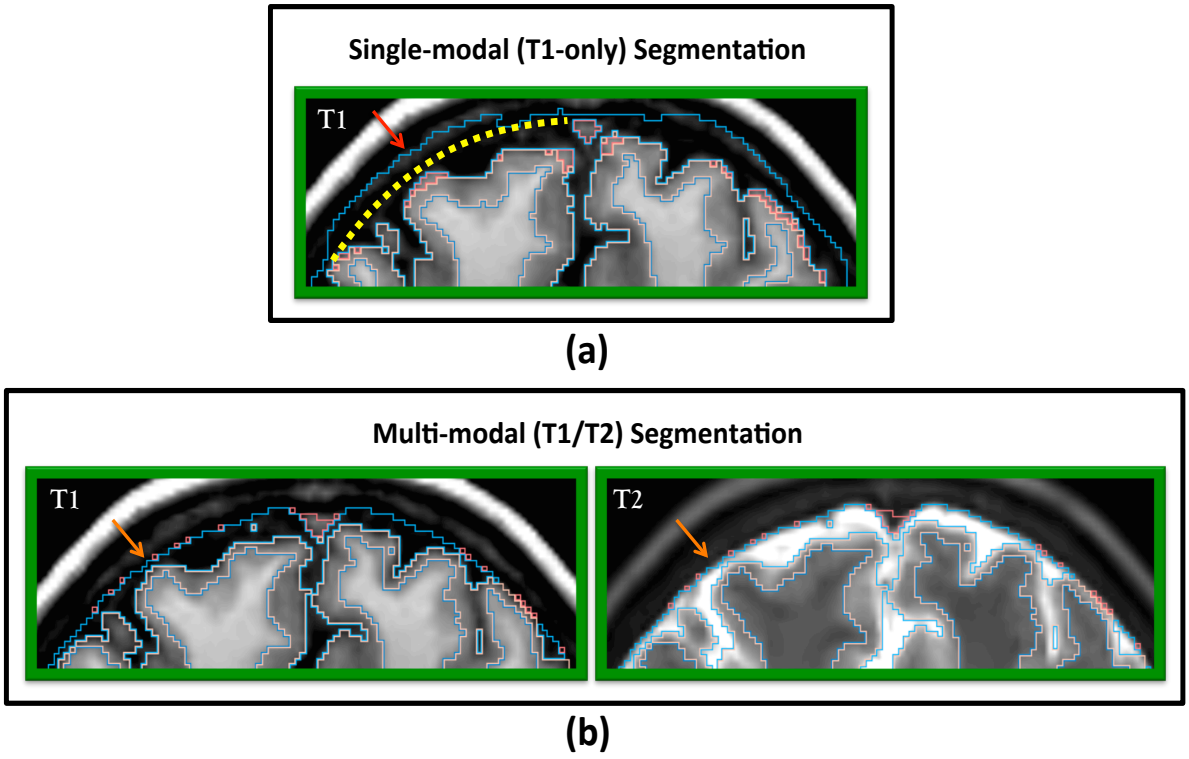


Figure 3.3: Visual comparison of segmentation results on a sample PREDICT-HD MR dataset (a) Single-modal segmentation of CSF using only T1-weighted modality. Segmentation results are shown in blue boundaries while the true anatomical boundary is defined in yellow based on the contrast provided by T2-weighted modality. (b) Multi-modal segmentation of CSF using both T1/T2-wegihted scans shows enhanced delineation of CSF as the segmentation results conform with true anatomical boundaries.

We also used the BrainWeb dataset, as described in section 2.3.1, to investigate the benefits of multi-modal segmentation quantitatively. It is demonstrated in Figure (3.4). *Red* color shows the results when a single-modal segmentation using only $T1$ MR modality is run, while *Green* color shows the results when two $T1$ and $T2$ MR modalities are used for the segmentation process while both modalities have the same isotropic 1 mm^3 resolution. Green lines show adding second modality has enhanced the segmentation results for CSF. Also, adding a second modality can enhance segmentation in higher levels of noise for WM and GM.

However, this does not hold when the additional multi-modal information comes from a lower spatial resolution. It is demonstrated again in Figure (3.4) by *blue* color when a multi-modal experiment is run using two $T1$ and $T2$ MR modalities with different spatial resolutions, where $T1$ is provided with isotropic 1 mm^3 resolution, but $T2$ has $1 \times 1 \times 3$ mm^3 voxel sizes.

Figure (3.4) demonstrates that adding information from a second modality with significantly lower resolution decreases the quality of segmentation results. It is expected because misalignment between voxel sizes and voxel locations increases the *partial volume effect* (PVE) at tissue boundaries, that means more spatial samples contain a mixture of different tissue types. Using these voxels to initialize EM parameters and to train $k - NN$ algorithm can adversely affect the performance of classifier and results in less accurate and less robust tissue classification.

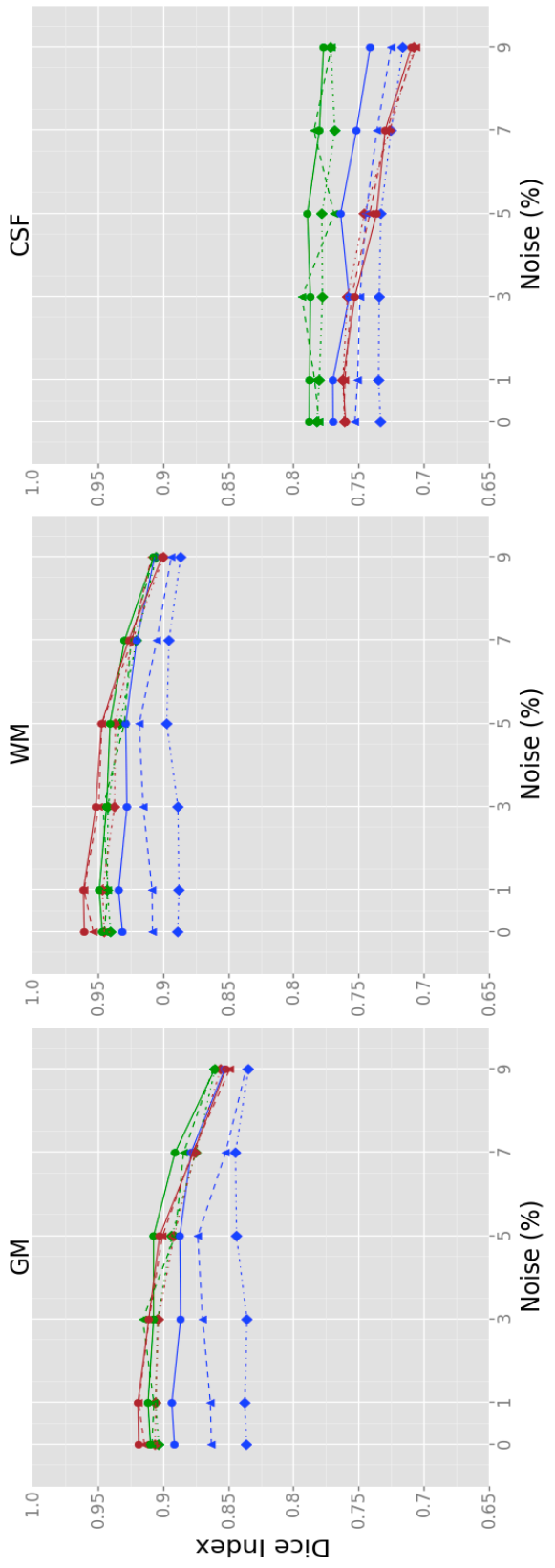


Figure 3.4: Blue lines show less accurate and less robust classification results when second modality ($T2$: $1 \times 1 \times 3\ mm^3$) has lower resolution than the first modality ($T1$: $1 \times 1 \times 1\ mm^3$).

3.3 Multi-modal Partial Volume Effect

Due to the finite spatial resolution of imaging device, a single image voxel may contain of several tissue types. This phenomenon, termed as partial volume effect (PVE), complicates the segmentation process, and, due to the complexity of human brain anatomy, dealing with the PVE issue is an important factor for accurate brain structure classification.

Figure (3.5) shows a schematic explanation of partial volume effect in a simplified $1D$ demonstration, where $1D$ MRI signals from three different image modalities (T1, T2 and DWI) are demonstrated near actual anatomical boundaries of the Gray matter and the White matter in brain.

The MRI works by detecting the magnetic particles in atoms within cells and sending electromagnetic pulses at different rates and strength through the body. Then, different types of matter give off different levels of energy when they are placed in a magnetic field. These energy signals, collected by the electromagnetic receivers, are quantized and sampled into discrete voxel segments in an output image scan. The different signal levels between different tissue types cause a different contrast for each matter in the output image. In an MRI scanning session, several different scans are run. Depending on how they measure the relaxation of magnetic particles, a variety of modality images with different contrasts are acquired. For example, in a T1 scan, gray matter gives off a low signal and looks darker than the white matter that gives off a higher signal and looks pale gray, but this is inverse in T2 and DWI scans.

Although tissue regions usually form coherent shapes with clear anatomical

borders, the representation of tissue boundaries in discretized image scan is not perfectly in agreement with real anatomy because of error in image discretization. In fact, biology cannot conform in the resolution levels that scanners produce.

Therefore, due to the hardware limitations and the presence of noise, the measurement of MRI signals is not perfect, and the collected signals are quantized in several intermediate levels in the boundary regions. This causes an inherent partial volume effect at tissue boundaries even in a single modality acquisition. The quantized signals are then mapped in discretized segments of spatial domain. A value is assigned to each voxel of image based on the average of all samples taken within the spatial range of that voxel. If the image has a low spatial resolution with large voxel sizes, it is most likely that many voxels are placed in two different anatomical regions, so their value reflects a partial volume composition of more than one tissue type.

Partial volume composition affects even a larger number of spatial samples when the multi-modal information comes from the modality scans having different resolutions and origins. This is shown in Figure (3.5) using a 1D demonstration. The MRI signals are sampled in 1 mm intervals. When only two T1 and T2 modalities with comparable resolutions are used, there are more pure spatial samples since only two samples are affected by partial volume composition (specified in a red box). However, adding new modality information from a low-resolution scan introduces the partial volume effect to a larger number of spatial samples. The affected samples are specified in a green box and represented by a question mark as they reflect more than one tissue type.

A careful consideration should be made to only use those spatial samples that are not affected by partial volume composition, termed as *pure samples*, for initializing/training of the classification modules.

3.4 Methods

Classification accuracy can be improved by excluding the voxels affected by the partial volume composition, termed as *mixed samples* or *nonpure samples*, from the initialization of EM algorithm and the training of $k-NN$ classifier. Figure (3.6) demonstrates the scatter maps of pure samples versus the mixed samples in a coronal view of two Gray matter and White matter tissue regions.

In this section, we develop a new approach to identify the pure samples by computing a binary mask, called *pure plugs mask*, that excludes all the mixed samples from the initialization/training of classification methods.

3.4.1 Computing Pure Plugs Mask

To compute a pure plugs mask, we should avoid the PVE inherent at tissue boundaries as well as the PVE caused by information inconsistency through multiple modality channels.

3.4.1.1 Avoid Inherent PVE

All the voxels located at tissue boundaries may include a mixture of both tissue regions since the biology cannot conform in the resolution levels that scanners produce. This causes an inherent partial volume effects that can be avoided by

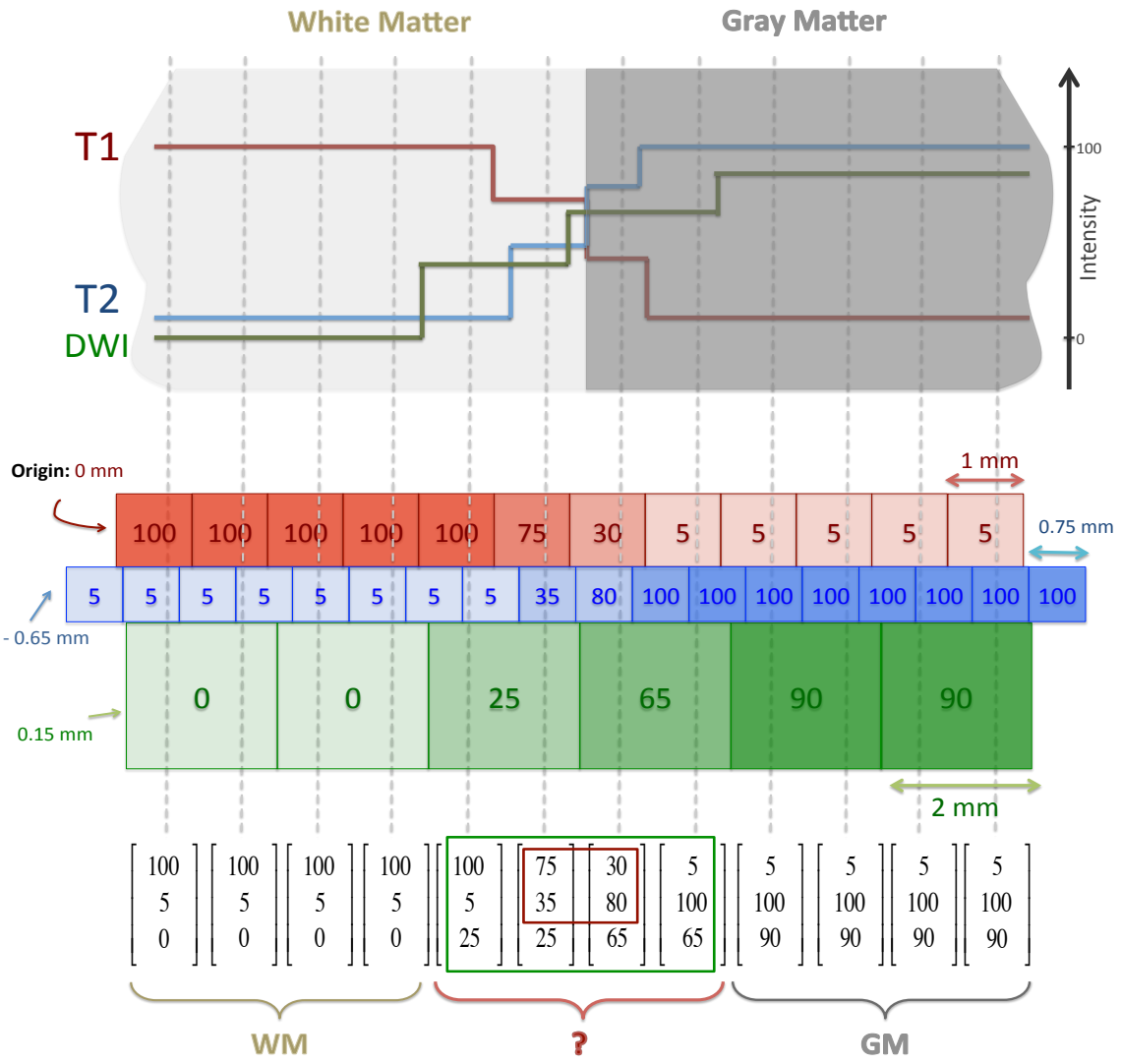


Figure 3.5: $1D$ demonstration of partial volume effect (PVE) when MRI signals of different modalities are sampled to discredited segments with different spatial resolutions and origins. Question mark represent the sample values that reflect a partial volume composition of more than one tissue type.

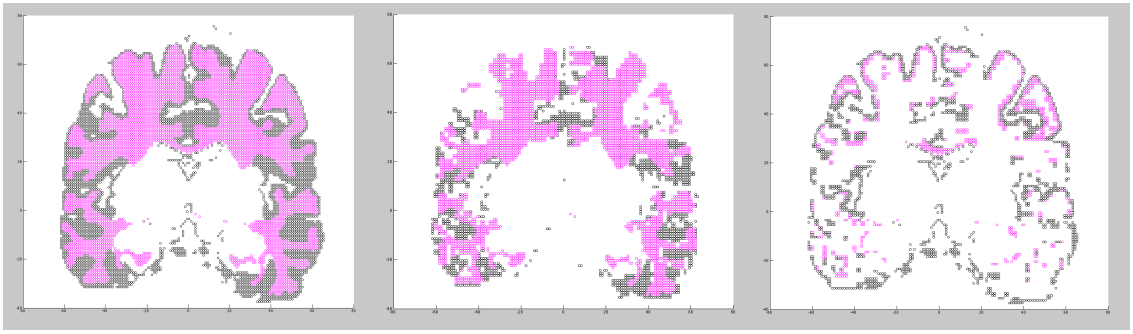


Figure 3.6: (a) a coronal view of a scatter point map of two Gray matter (black) and White matter (pink) tissues. (b) only pure samples are shown. (c) only mixed samples are shown.

excluding all tissue boundaries from the pure plugs mask. For this purpose, anatomical edges are detected from the input modality with the finest resolution (usually a $T1$-weighted scan) using a Canny edge detector [74] as shown by Figure (3.7).

3.4.1.2 Avoid Multi-modal PVE

As demonstrated by Figure (3.5), partial volume composition affects a larger number of spatial samples when the multi-modal information comes from multiple modality scans with different voxel resolutions. To deal with this issue, only pure samples should be included in the classification process. To find the pure samples, *pure plugs* should be computed.

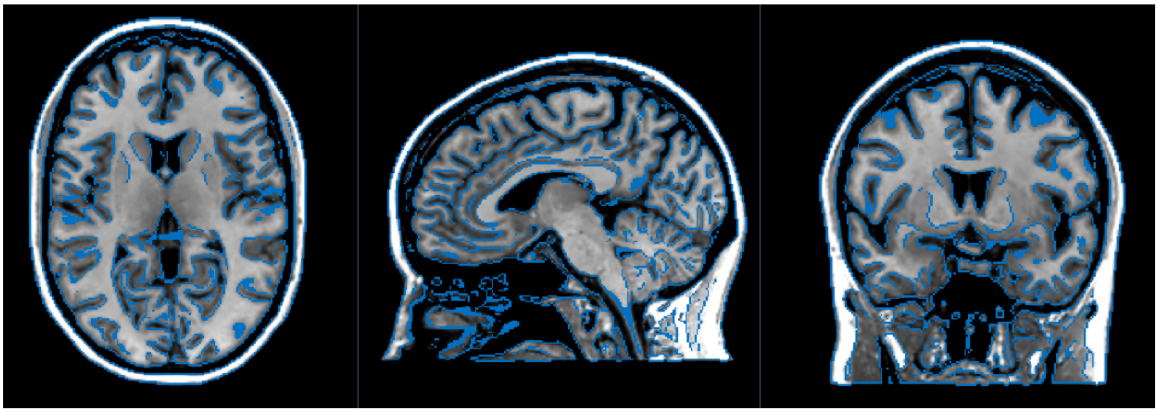


Figure 3.7: Anatomical tissue boundaries are detected using a Canny edge detector. Detected edges should be excluded from the pure plugs mask to avoid inherent PVE in tissue boundaries.

3.4.1.2.1 What Is a Pure Plug?

A *pure plug* is defined as a region (plug), where all multi-modal images have consistent information within the entire range of that region. The size of each plug region is then decided based on the lowest spatial resolution at each direction within all the input modality scans.

This definition raises an immediate question: how do we define the consistency of multi-modal information within a plug region? This is discussed in following section by defining the *pure plugs integrity metric*.

However, before analyzing the information from multi-modal channels, all the input images should be normalized to have the same dynamic range for their intensity values.

3.4.1.2.2 Standardize Intensity

An intensity transformation function, as shown in Figure (3.8), is applied to the intensity levels of each input image. All the input modality scans are scaled to have the same dynamic range based on the first and last 5 percentiles of their histogram. Also, low and high tails of data in the output image are trimmed to the constant bounds of 0 and 1. This transform function can be formulated as:

$$O(i) = \begin{cases} \alpha I(i) + b_0 & \text{if } \frac{-b_0}{\alpha} < I(i) < \frac{1-b_0}{\alpha} \\ 1 & \text{if } I(i) > \frac{1-b_0}{\alpha} \\ 0 & \text{if } I(i) < \frac{-b_0}{\alpha} \end{cases} \tag{3.1}$$

$$where \quad \alpha = \frac{0.9}{Q_I(95) - Q_I(5)} \ , \ b_0 = 0.95 - Q_I(95).\alpha$$

Where at index i, $O(i)$ is the output image intensity, and $I(i)$ is the input image intensity. Also, $Q_I(p)$ is the p^{th} quantile (percentile) of the input intensity range.

3.4.1.2.3 Pure Plugs Integrity Metric

To decide whether a plug is considered pure, several uniformly distributed spatial sample points should be taken within the entire range of the plug area. Then, each sample point is represented in an N-dimensional intensity space based on the normalized intensity values from N input modality channels.

Figure (3.9) shows a joint image histogram for two input modality, $T1$ and $T2$-weighted, images. The picture shows the distribution of points in different colors

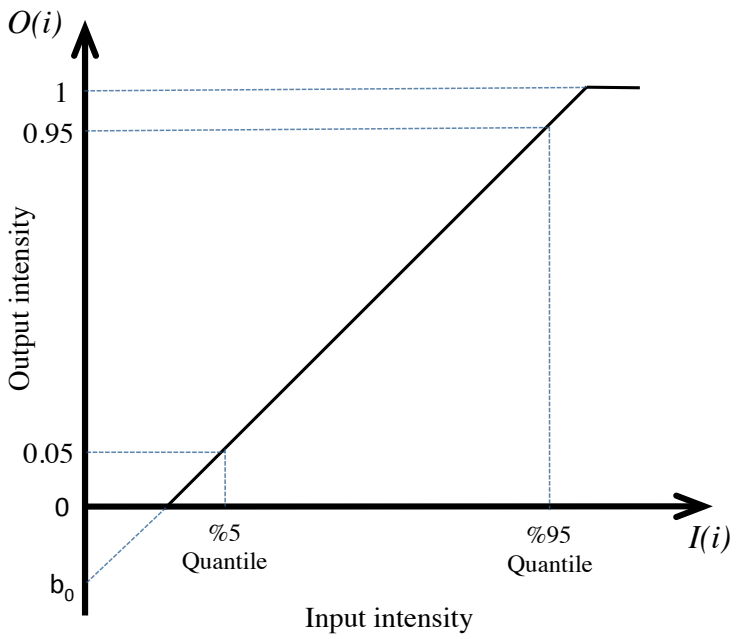


Figure 3.8: Intensity transform function applied to each input modality image to make all the input images have the same intensity dynamic range.

for background and four tissue regions (White Matter (WM), Gray Matter (GM), Cerebrospinal Fluid (CSF), and Venous Blood (VB)). It is important to notice that the points are scattered non-uniformly in different directions for a single tissue region. The reason is that the input modality scans have different variances for the intensity values within a tissue region.

This means that the spatial sample points taken within a plug region may not be uniformly distributed in the intensity space even if all belong to a single tissue region. Hence, a metric should be defined to decide about the integrity of sample points based on following criteria:

- Detects the possible outliers in intensity distribution of the sample points.

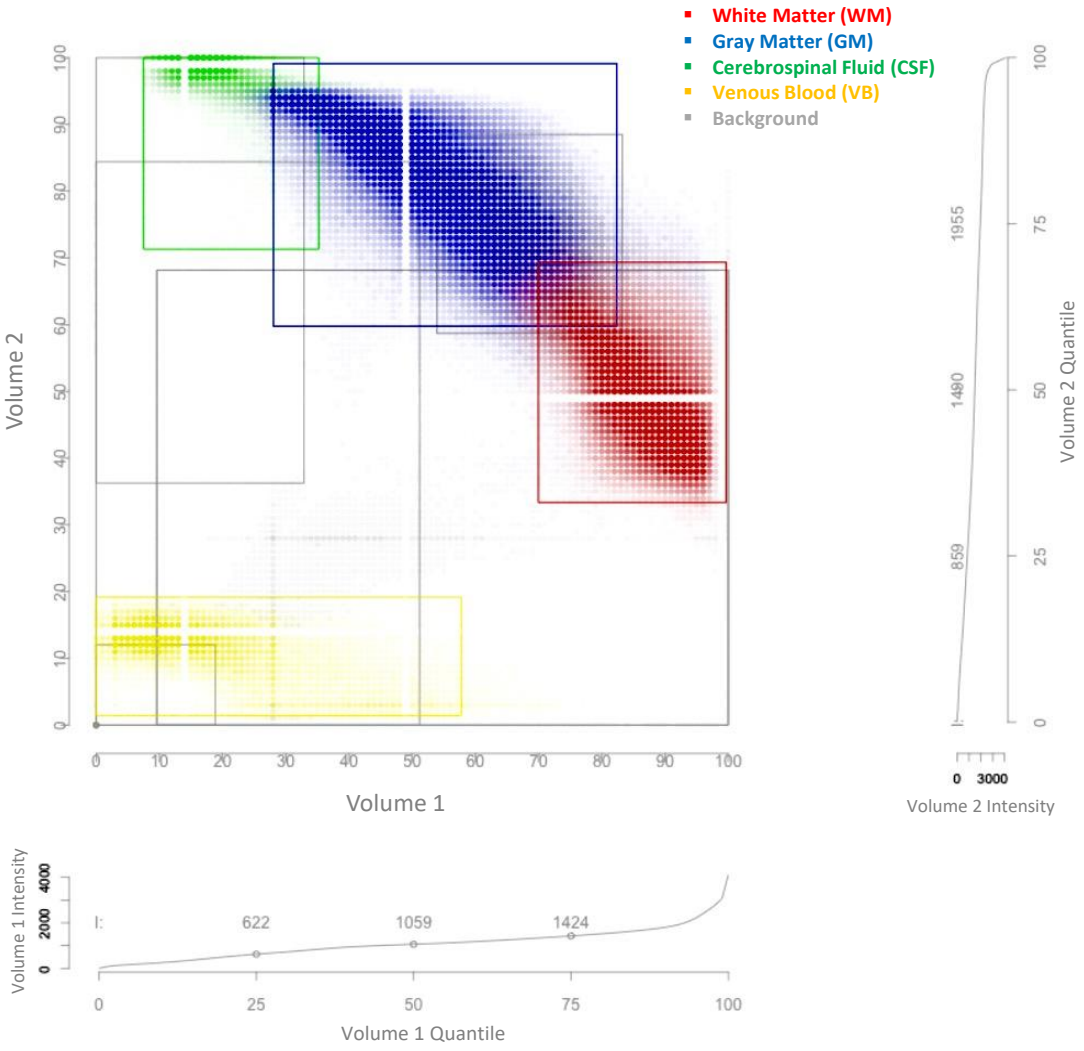


Figure 3.9: Joint image histogram for two modality images. Volume 1 is a T1-weighted image, and volume 2 is a T2-weighted scan from the same subject. The distribution of points is shown in different colors for background (grey) and four tissue regions: White Matter (red), Gray Matter (blue), Cerebrospinal Fluid (green), and Venous Blood (yellow). A box is plotted around the mean of each region with the size of 4 standard deviation at each direction.

- Considers how far the sample points are scattered from each other.

The shape and size of multivariate sample points can be quantified by a covariance matrix. *Mahalanobis distance* is a well-known distance measure that takes into account the covariance matrix, so it is used as the basis for multivariate outlier detection [75]. For an N-dimensional sample $\bar{X}_k = (x_1, x_2, ..., x_N)^T$ in a set of observations, the Mahalanobis distance is defined as:

$$D_M(\bar{X}_k) = \sqrt{(\bar{X}_k - \bar{\mu})^T S^{-1} (\bar{X}_k - \bar{\mu})} \tag{3.2}$$

Where S is the covariance matrix, and $\bar{\mu} = (\mu_1, \mu_2, ..., \mu_N)^T$ is the mean of all observations. In fact, the Mahalanobis distance is a Euclidean distance that considers the covariance of data by down-weighting the axis with higher covariance. Then, a sample $\bar{X}_n$ is considered as outlier, if $D_M(\bar{X}_n) > \alpha.Mean(D_M)$, where α can tune the threshold value.

Although the Mahalanobis distance is useful to detect the outliers, it is not a good integrity metric for finding the pure plugs, as it is scale invariant. That means that Mahalanobis distance does not consider how far the sample points are scattered. Figure (3.10) shows this using an example. To consider the scatter of points, a *Euclidean distance* can be used to show how dense all the sample points are scattered within the range of intensity space ($0 < i < 1$).

Therefore, a combination of both *Mahalanobis* and *Euclidean* distance measures can be used to define the integrity metric, such that:

- Mahalanobis distance considers the covariance of multivariate sample points

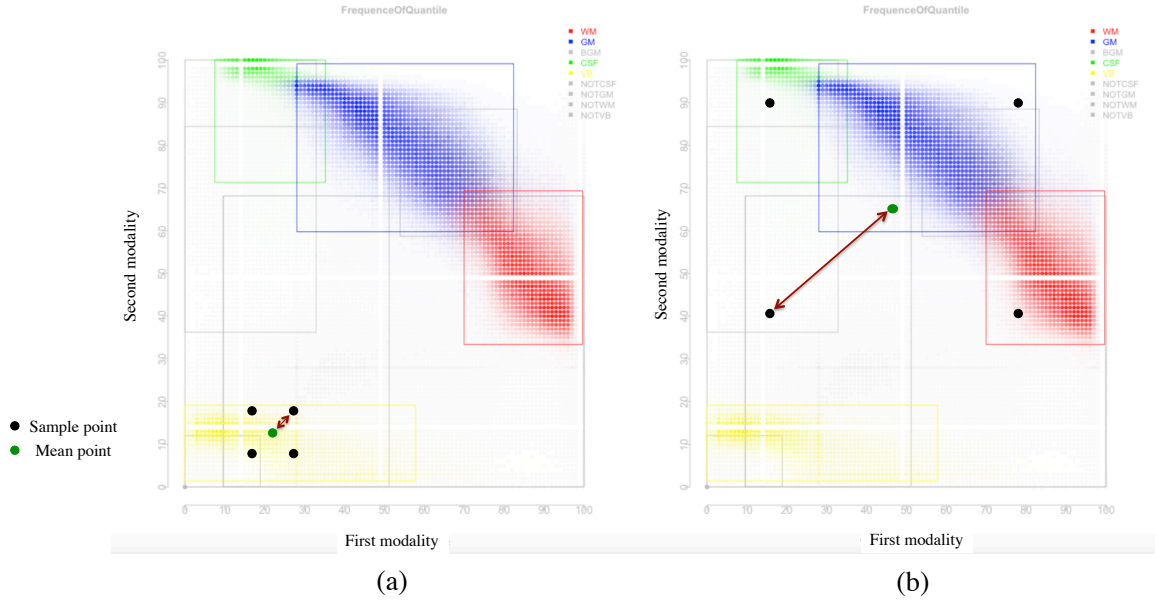


Figure 3.10: Mahalanobis distance is scale invariant and does not consider the range of scattered points. (a) The Mahalanobis distance for all points is 1.22, and all points belong to a single tissue region. (b) The Mahalanobis distance for all points is 1.22, but points belong to four different tissue regions.

distribution to detect the outliers.

- Euclidean distance considers the density of sample points distribution within an intensity range.

The introduced integrity metric then combines both characteristics by using the normalized Mahalanobis distance to weight the Euclidean distance for each sample point taken within the plug region. The integrity metric, called **Mahalanobis-weighted Euclidean distance**, is then defined as:

$$D(i) = \frac{M_D(i)}{Max(M_D)} \times E_D(i) \tag{3.3}$$

Where, for the i^{th} sample point, $D(i)$ is the introduced integrity distance metric, $M_D(i)$ is the Mahalanobis distance, $E_D(i)$ is the Euclidean distance to the mean, and $Max(M_D)$ is the maximum Mahalanobis distance between all observations.

The computed integrity metric distance for all the sample points in Figure (3.10 a) is 0.07, that is significantly smaller compared to the computed distance metric of 0.42 for the sample points in Figure (3.10 b). Thus, choosing a threshold value of 0.1 can indicate that the samples in Figure (3.10 a) belong to a pure plug region, while the sample in Figure (3.10 b) do not belong to a pure plug.

Based on the introduced distance metric in equation (3.3), following steps should be taken at each plug region to decide whether it is pure:

1) Several spatial sample points are taken within each plug area, such that the points are uniformly scattered over the entire region.

2) All samples are transformed to an N-dimensional space based on their intensity

values from N input modality channels.

3) The integrity distance measure, as defined in equation (3.3), is computed for each sample point considering the covariance and density of distribution.

4) The plug is then considered to be pure, if all the integrity metric values are less than a threshold value that is defined as a value in the normalized intensity range (between 0 and 1).

The threshold value can be tuned based on the dataset. Here, we empirically chose a value of 0.2. Figure (3.11) shows a test pure plugs mask created based on 3 modalities from a test dataset listed in Table (3.1). This dataset was used as part of multi-site international PREDICT-HD project. The third modality is an isotropic diffusion-weighted information (IDWI) image, that is the geometric mean of all diffusion images in an input DWI dataset. This image has voxel sizes 8 times larger than typical structural MR modalities.

Table 3.1: Test dataset to create a test pure plugs mask.

Scan	Site	MR vendor	Field strength	Collected modalities
0140_52100	Site_024 (University of Iowa)	SIEMENS TrioTim	3.0	T1: $1 \times 1 \times 1\ mm^3$ T2: $1 \times 1 \times 1\ mm^3$ IDWI: $2 \times 2 \times 2\ mm^3$

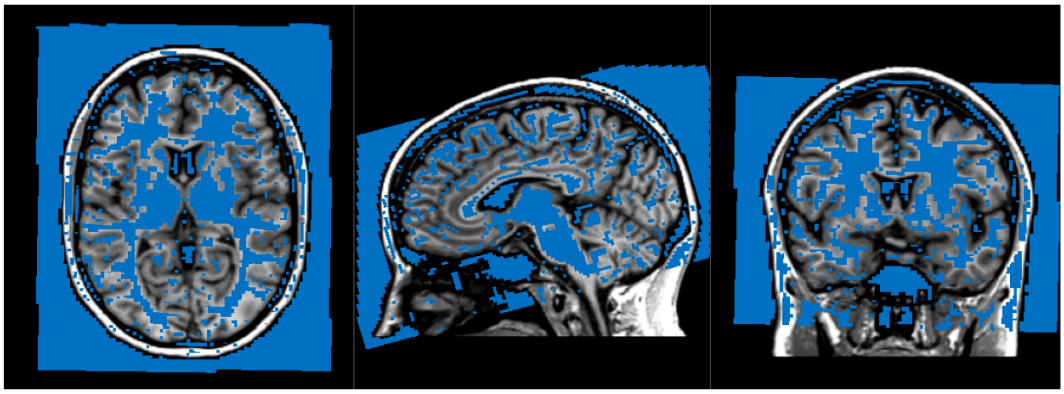


Figure 3.11: A test pure plugs mask created from 3 modality scans from a subject test listed in table (3.1).

3.5 Experimental Methods and Results

In this section, we evaluate the segmentation enhancement when only pure samples, detected by computing a pure plugs mask, are used in initialization/training of classification methods.

3.5.1 2D Phantom Analysis

Figure (3.12) illustrates two 2D phantom images, representing two modalities, and their corresponding baseline label map to run a sensitivity analysis that evaluates the accuracy of segmentation over the downsampling factor of second modality.

Both images have the size of 128×128 with spacing of $1 \times 1\ mm^2$. Figure (3.12 a) represents the first modality with two tissue types (A and B) in addition to the background, where tissue type A is, in fact, a composition of two anatomical tissues that are not distinguishable by the first modality phantom as they are represented by the same intensity value. The values in parentheses show the assigned intensity value

to each tissue region. Figure (3.12 b) represents the second modality with three tissue types in addition to the background. The second modality phantom shows enough contrast to distinguish between tissue types A and C.

For the sensitivity analysis, we downsampled the second modality by different downsampling factors (1 to 10), and at each step the pure plugs mask was generated from the first modality (in high resolution) and the downsampled second modality (in low resolution). As an example, Figure (3.13 a) shows the second modality downsampled by a factor of 10. The corresponding pure plugs mask is shown overlaid on the image in Figure (3.13 b).

At each downsampling step, the segmentation error rate was computed using each of following experiments:

- A high-resolution estimation of second modality was reconstructed from the downsampled low-resolution image using a nearest neighbor interpolation, and a class label was assigned to each pixel based on its intensity value.
- A fuzzy K-Nearest Neighbor classifier was used to generate the segmentation label map when a random set of samples were used to train the $k-NN$ classifier.
- A fuzzy K-Nearest Neighbor classifier was used to generate the segmentation label map when only pure samples (within the pure plugs mask) were used to train the $k-NN$ classifier.

Figure (3.14) compares the results from above three experiments. The segmentation error rate was calculated as the percentage of misclassified pixels to the total number of pixels when the segmentation results were compared to the baseline

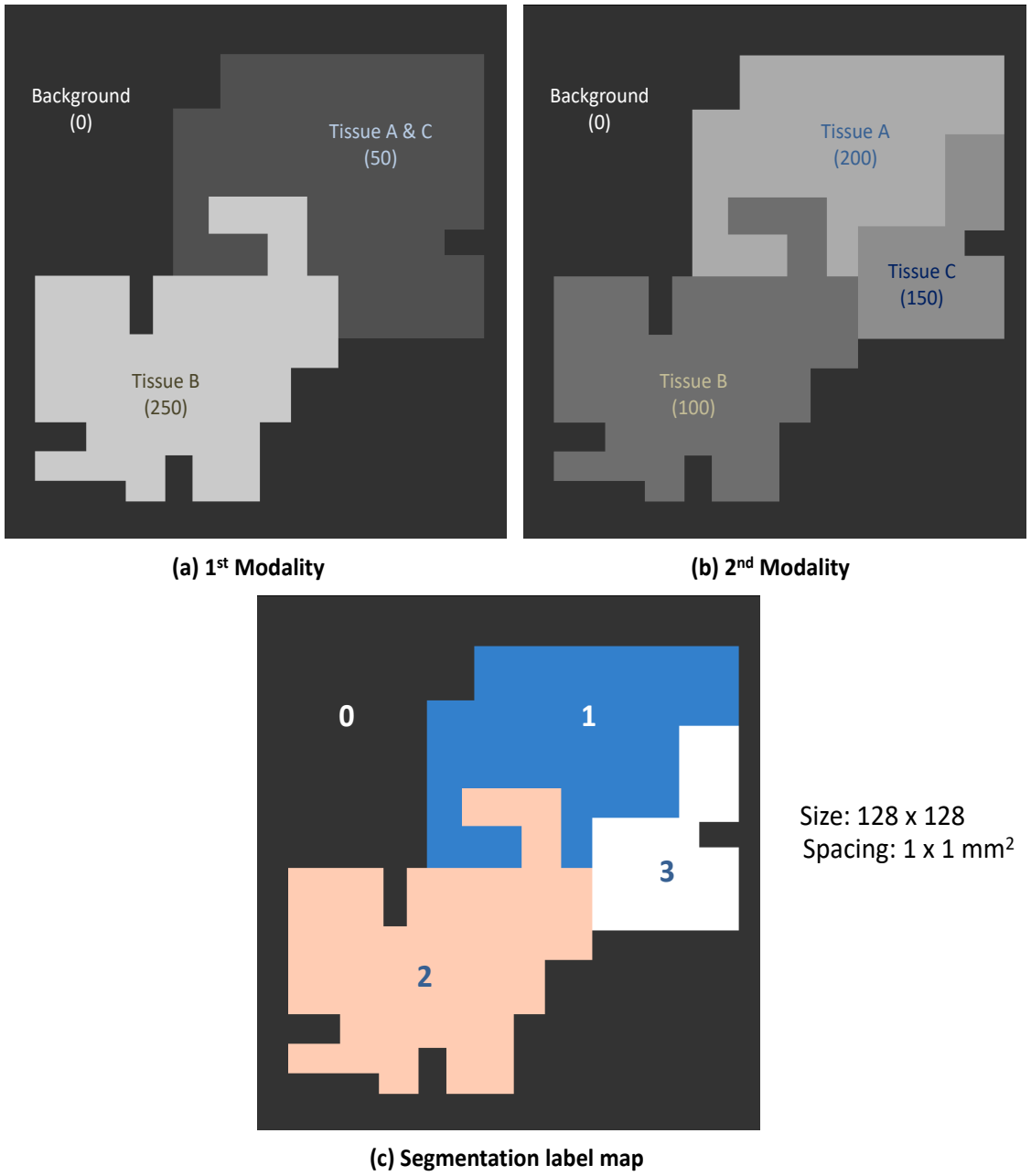


Figure 3.12: Two 2D phantom images representing two modalities and their corresponding baseline label map.

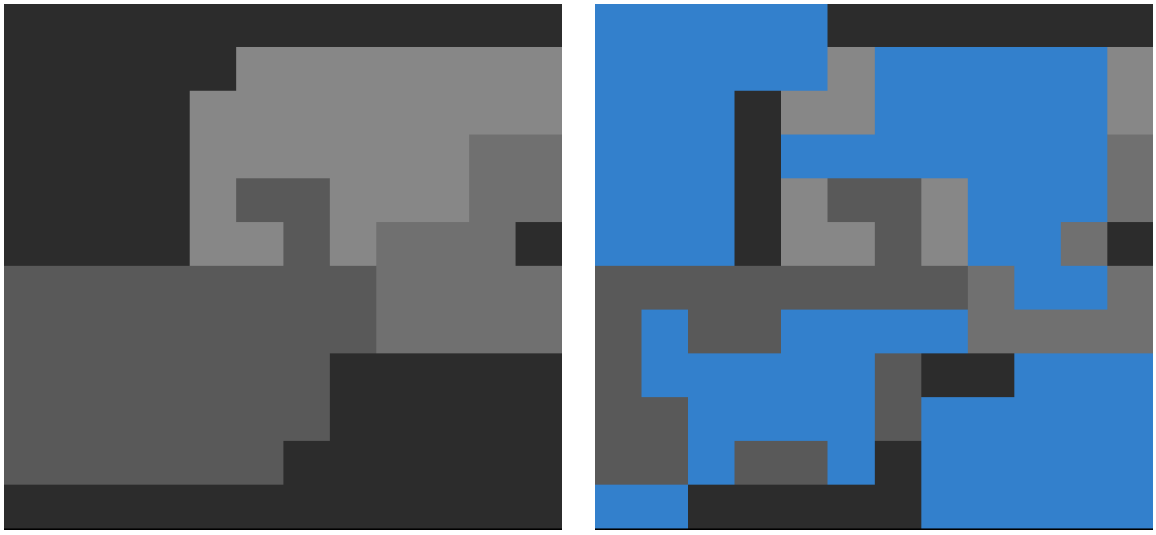


Figure 3.13: (a) The second modality downsampled by a factor of 10. (b) The corresponding pure plugs mask (in blue) overlaid on the downsampled image.

label map illustrated in Figure (3.12 c).

The results show that classification approaches perform better than simple upsampling by nearest-neighbor interpolation. Although both classification approaches show increased error rate with higher downsampling factor, the developed novel approach, using pure plugs mask, reduces the segmentation error rate from about 10% to less than 8% when only pure samples were use for the training of the classifier. The developed algorithm cannot help much in addressing the partial volume effects in boundaries between tissues A and C because the boundary information only comes from the low-resolution second modality data. However, the developed method helps to reduce misclassification in boundaries between tissues A and B, and between tissues C and B, as well as the boundaries between the background and each tissue

region. Choosing only pure samples for training increases the distinguishability between different tissue classes in the feature space as unillustrated by Figure (3.15). This leads to more accurate classification of mixture boundary pixels that are affected by partial volume composition.

3.5.2 Qualitative Evaluation

3D Slicer [64] was used to visually compare the previous segmentation results to the results of developed approach when only pure samples are used to initialize/train the classification methods. Qualitative investigation was done using a sample dataset that was selected from a GE Signa 1.5 Tesla scan protocol, where $T1$ resolution is $1 \times 1 \times 1\ mm^3$, but $T2$ voxel size is $1.015 \times 1.015 \times 3\ mm^3$. This protocol was used as part of multi-site international PREDICT-HD [1] project.

Figure (3.16) shows the qualitative results. It shows that incorporating pure samples in the classification process can help detecting subtle tissue regions that were missed when the second modality has a lower spatial resolution. Also, new approach enhances the segmentation accuracy in complex tissue boundaries.

3.5.3 Quantitative Evaluation

Developed enhancements were evaluated quantitatively using BrainWeb database as described in section 2.3.1.

Figure (3.17) reports the results of four experiments. *Green* color shows the results when two modalities are used ($T1$ and $T2$), and both images have the same isotropic 1 mm^3 voxel lattice. *Blue* color shows the results when two modalities with

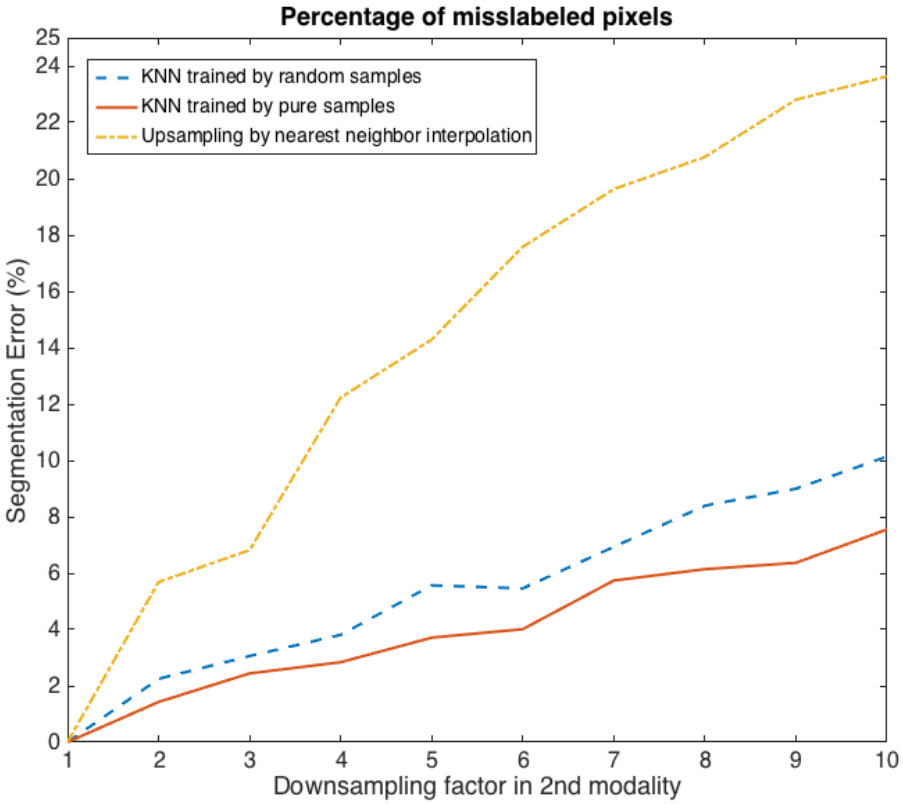


Figure 3.14: The segmentation error rate is computed as the percentage of misclassified pixels to the total number of pixels when the baseline label map is compared with the results of multi-modal segmentation from three different experiments: (1) The low-resolution modality was upsampled using a nearest neighbor interpolation. (2) A fuzzy K-Nearest Neighbor classification was run when a random set of samples were used to train the classifier. (3) A fuzzy K-Nearest Neighbor classification was run when only pure samples were used to train the classifier.

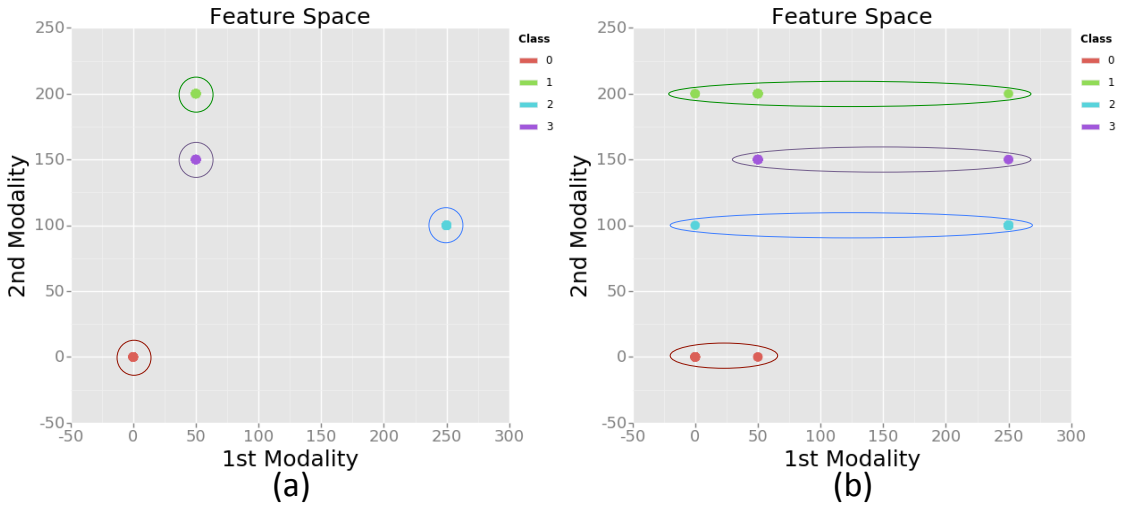
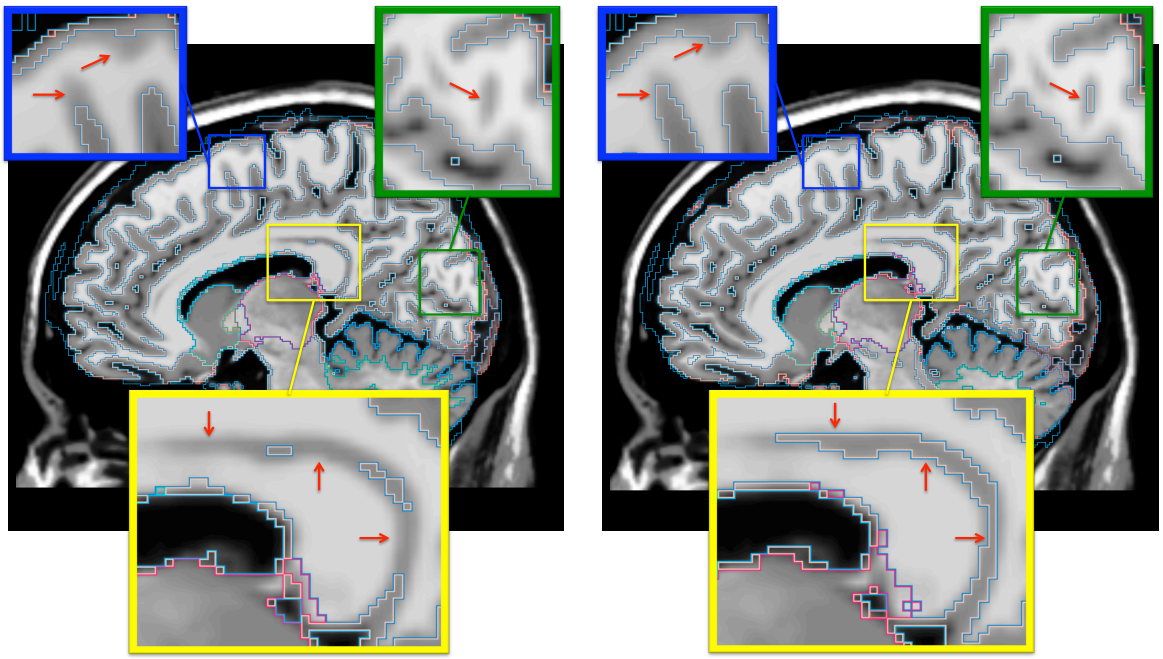


Figure 3.15: Training samples and their corresponding class labels in the feature space. (a) Only pure samples are using for training. (b) Both pure and non-pure samples are used for training.

different resolutions are used. $T1$ is $1 \times 1 \times 1$ mm^3, but $T2$ has the voxel size of $1 \times 1 \times 3$ mm^3. In both cases, solid lines show the segmentation results when pure plugs mask is NOT used, while dashed lines show the segmentation results when pure plugs are computed and incorporated in the segmentation process.

We also investigated the regional benefits of using pure samples in segmentation by showing the results of the gray matter (GM) and white matter (WM) in each of four different lobes of brain (frontal, occipital, temporal and parietal (Figure 3.18)). Figure (3.19) shows the results for GM and Figure (3.20) presents the results for WM.



(a) (b)

Figure 3.16: (a) Segmentation results when pure plugs mask is not incorporated in classification process. (b) Results when only pure samples are used to initialize/train the classification methods. Incorporating pure samples in the classification process enhances segmentation in tissue boundaries and can help better delineation of subtle tissue regions.

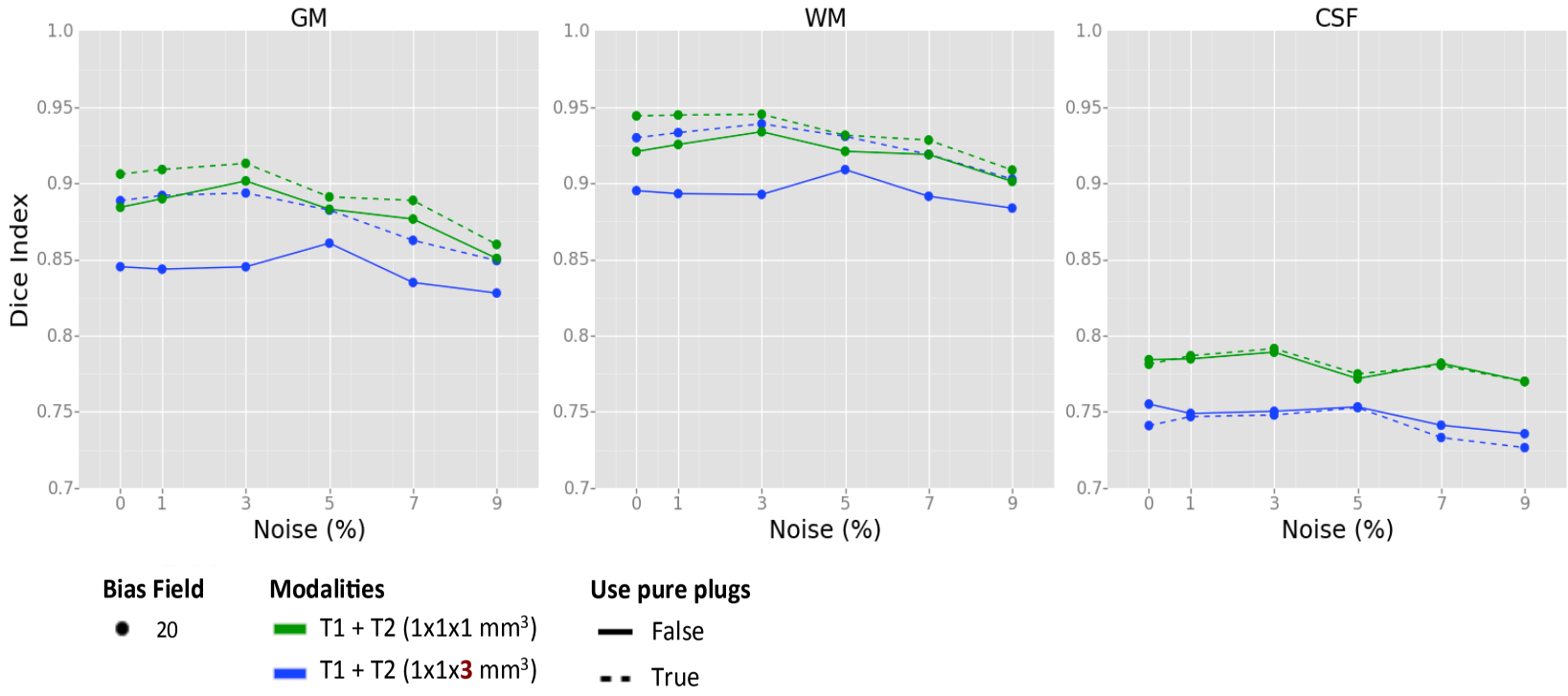


Figure 3.17: Green lines show the results when the multi-modal segmentation is run using two modalities with the same isotropic 1 mm^3 resolution. Blue lines show the results when $T2$ modality has a lower spatial resolution as $1 \times 1 \times 3$ mm^3. Dashed lines show the improved results when only pure samples are involved in the classification process. This improvement is more significant when the second modality has lower spatial resolution as indicated by blue lines.

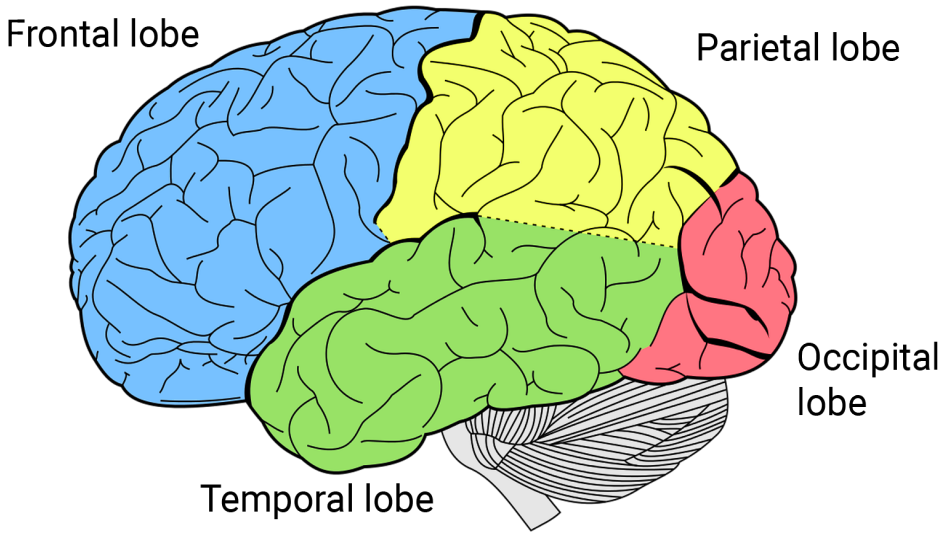


Figure 3.18: Human brain diagram showing four different lobes of brain *(Image: Public domain).*

3.6 Discussion and Conclusions

Both qualitative and quantitative results proved that incorporating pure plugs in the classification methods always enhances the segmentation quality for gray matter (GM) and white matter (WM), such that the accuracy of multi-modal classification with low-resolution T2 and using pure plugs mask approaches the quality of segmentation when all multi-modal scans are provided in high-resolution.

Also, the regional results for gray matter and white matter in each lobe of brain conforms with the results computed within the whole cerebrum. In occipital lobe, although incorporating pure plugs enhances the segmentation results, the quality of low-resolution multi-modal segmentation is still lower than the segmentation quality when both modalities are provided in high spatial resolution because occipital

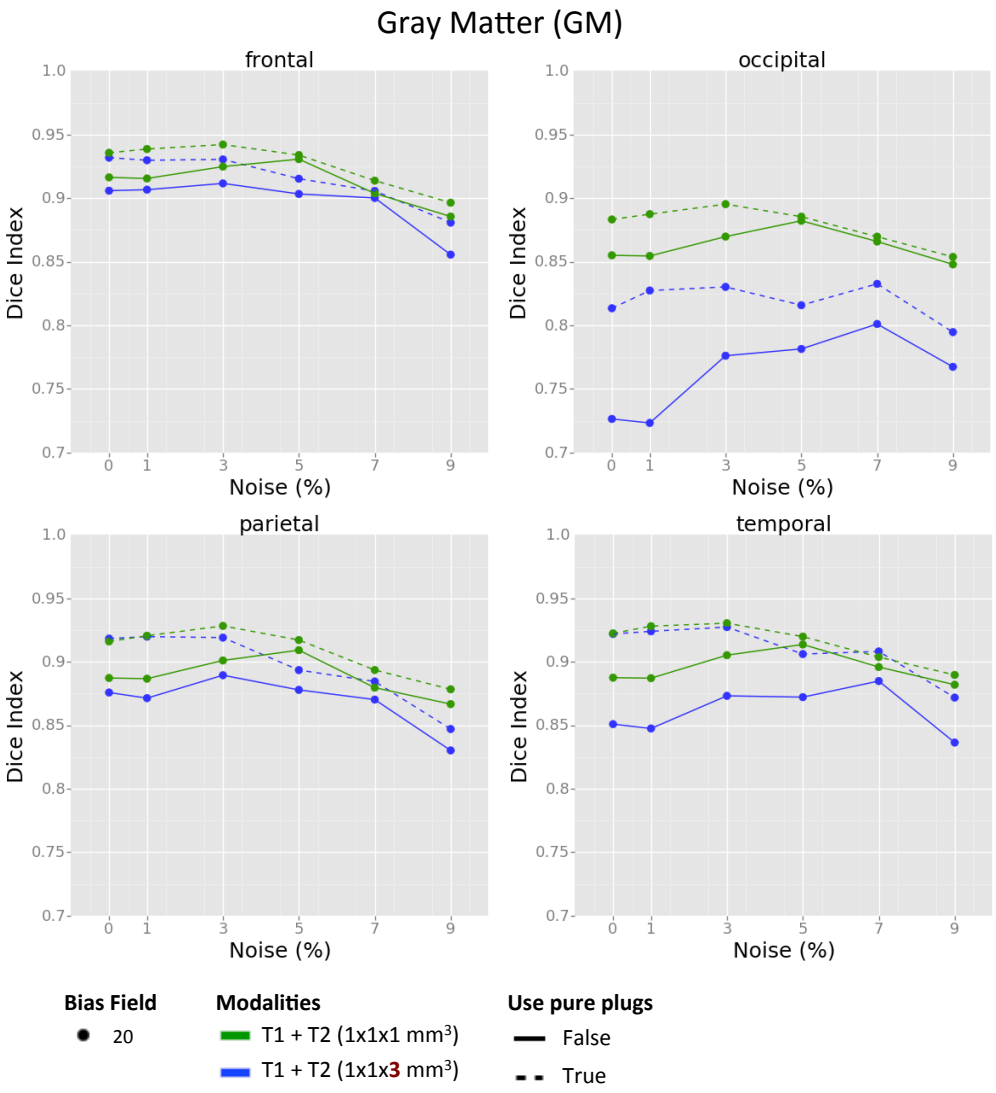


Figure 3.19: Regional benefits of using pure samples in segmentation of gray matter (GM) in in each of four different lobes of brain (frontal, occipital, temporal and parietal). Green lines show the results when the multi-modal segmentation is run using two modalities with the same isotropic 1 mm^3 resolution. Blue lines show the results when $T2$ modality has a lower spatial resolution as $1 \times 1 \times 3\ mm^3$. Dashed lines show the improved results when only pure samples are involved in the classification process.

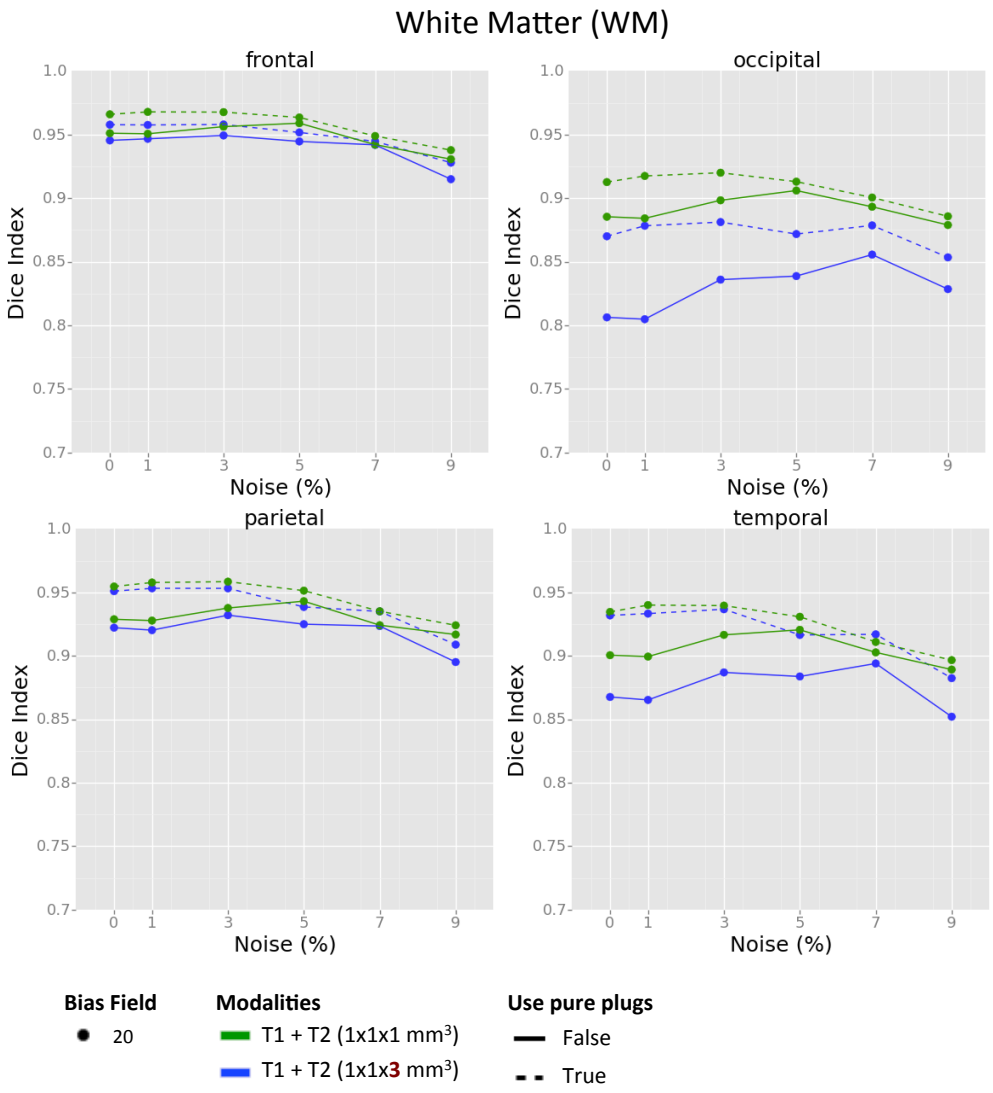


Figure 3.20: Regional benefits of using pure samples in segmentation of white matter (WM) in in each of four different lobes of brain (frontal, occipital, temporal and parietal). Green lines show the results when the multi-modal segmentation is run using two modalities with the same isotropic 1 mm^3 resolution. Blue lines show the results when $T2$ modality has a lower spatial resolution as $1 \times 1 \times 3\ mm^3$. Dashed lines show the improved results when only pure samples are involved in the classification process.

lobe is a relatively smaller region where many voxels are affected by partial volume composition.

For the CSF, incorporating pure plugs does not show significant improvement to the quality of segmentation results. This happens because the introduced method in this chapter is expected to show significant improvements in complicated tissue boundaries with high partial volume effects like the regions where WM and GM are highly interleaved together. However, CSF is mostly included in large pool areas of ventricles that are less affected by partial volume composition.

This chapter showed that by dealing with partial volume effects (PVE), multispectral information can be combined from both high-resolution and low-resolution modalities to enhance the segmentation quality of a subject dataset, while naïvely adding the low-resolution data would adversely affect the segmentation results.

The new developments introduced in this chapter directly improved segmentation quality for the longitudinal, multi-site, international PREDICT-HD study [1], as now it is possible to use the low-resolution T2-weighted modalities that were acquired during the first 7 years of this study to enhance the segmentation results previously generated using only T1-weighted scans.

The developed method may also help to reduce the cost and the scanning time in future clinical trials as by dealing with PVE algorithmically, we can collect only the first modality in a high resolution and use the other modalities acquired in lower spatial resolutions. Decreasing the scanning time may have other benefits like reducing the subject burden for being in the scanner.

CHAPTER 4
SUPER-RESOLUTION RECONSTRUCTION OF LOW-RESOLUTION DIFFUSION-WEIGHTED IMAGING DATA USING *A PRIORI* KNOWLEDGE OF ANATOMICAL STRUCTURES

4.1 Introduction

Diffusion-weighted imaging (DWI) is a key imaging modality that enables non-invasive and in-vivo investigation and characterization of brain white matter (WM) architecture and microstructure and is widely applied in neurological applications.

DWI is, however, strongly limited by its relatively low spatial resolution. The resolution of DWI data that is compatible with the time constraints of clinical research is $2 \times 2 \times 2\ mm^3$. A DWI voxel volume is approximately 8 times larger than that of a typical structural MRI. Increasing the resolution of DWI acquisitions can allow investigation of novel fiber structures and will enable a more accurate assessment of brain connectivity by tracing small white matter fiber bundles. Also, high resolution images are critical to reduce partial volume effects. However, increasing the resolution is challenging in DWI. A DWI scan needs to be repeated 64 times for averaging to increase the resolution from $2 \times 2 \times 2\ mm^3$ to $1 \times 1 \times 1\ mm^3$ while keeping the similar signal-to-noise ratio [76]. It means that a 5-minute acquisition would become a 5 hour scan, that is not feasible.

To enhance the resolution, image post-processing methods are an alternative to hardware improvement. However, simply using the interpolation methods to increase the resolution causes results show blurry edges. The term of super-resolution

reconstruction (SRR) refers to considering image degradation process to estimate the latent high-resolution image from the input low-resolution scan.

To overcome the limitations of low-resolution DWI, some SRR methods have been recently developed. One group of methods require multiple low-resolution (LR) images to reconstruct a high-resolution (HR) image. Scherrer *et al.* [77] suggested to acquire multiple anisotropic orthogonal DWI scans and fuse them into a high resolution output. Ning *et al.* [78] combined the concept of compressed sensing and classical super-resolution to reconstruct high-resolution DWI from multiple sub-pixel-shifted thick-slice LR acquisitions with non-overlapping diffusion directions to reduce acquisition time. However, these types of methods are hampered for general applications because: first, a specially designed image acquisition method is needed to acquire multiple scans; second, the subject motion and eddy current effects in different scans could largely affect the final results. The other group of methods obtain HR data using a single LR image through a learning process or an intelligent regularization. Alexander *et al.* [79] proposed a method to exploit information from expensive high quality datasets and transfer them to enhance the images acquired from a more modest data acquisition. Their method attempts to learn mapping from LR to HR through training sets using patch-based image representation and random forest regression. Tarquino *et al.* [80] suggested a patch-based sparse representation approach to recover HR reconstruction using the coupled low and high resolution dictionaries. Although these methods do not require multiple LR acquisitions, they still need a separate high-resolution DWI training dataset. Shi *et al.* [81] proposed a

method for super-resolution reconstruction of a single LR scan by modeling degradation process and use of different regularization terms with no need of a high-resolution training dataset. However, the performance of their method still largely relies on the information contained in the original low-resolution image. As an example of edge-directed interpolation for resolution enhancement, Li *et al.* [82] have developed a multi-resolution covariance-based adoption method around edge pixels in the image. They estimated the HR covariance from the LR counterpart based on their geometric duality. The application of this method to DWI super-resolution has not been investigated. Also, the performance of their method relies on the information contained in the original LR image and the model employed to describe the relationship between HR pixels and LR pixels. As an example of using anatomical priors, Yendiki *et al.* [83] have developed an algorithm to reconstruct a white matter pathway jointly from a series of longitudinally acquired DW images. In their study, they generated a within-subject template from the T1 images of the subject at all time points. Then, they followed a global probabilistic tractography approach to present an unknown WM pathway in the space of the within-subject template and propagate to the native space of DWI at all time points to compute its posterior probability given the images. The application of their method is in longitudinal studies to reconstruct an unknown WM pathway.

Here, we develop a novel method for super-resolution reconstruction of a low-resolution input DWI image using the prior anatomical information extracted from other modality sources that are provided in higher spatial resolution. Particularly, we

incorporate some anatomical descriptions provided by high-resolution structural MR (e.g. T1/T2-wighted) scans into the super-resolution reconstruction of DWI image. Our method aims to increase the resolution of input low-resolution DWI to a higher spatial resolution, e.g., at a factor of 2. Our contribution is twofold: 1) Create a combined edge map from the structural MR scans in higher spatial resolution ($1\times1\times1$ mm^3); 2) Use the created edge map as discrete spatial weights to run super-resolution reconstruction in an edge-guided weighted total variation (WTV) method.

4.2 Mathematical Background

4.2.1 Regularized Recovery of Inverse Problems

Image reconstruction is the process of recovery of an ideal intensity image from its corrupted or indirect measurements. Such a process is considered as an inverse problem in science as it starts with the results (observations) and then calculates the causes. The observed data are usually related to the ideal unknown image through a "forward" transformation [84].

We consider the recovery of a continuously differentiable image $f : \Omega \to \mathbb{R}$ from its measurements b. Here, $\Omega \subset \{\mathbb{R}^n \mid n = 2 \quad or \quad 3\}$ is the spatial support of the image. The acquisition scheme is modeled by a linear operator $\mathcal{A}$, i.e.,

$$b = \mathcal{A}(f) + n \tag{4.1}$$

Where n is assumed to be a Gaussian distributed white noise with standard deviation of σ.

The recovery is ill posed in many practical applications as the operator $\mathcal{A}$ is

ill conditioned. One popular approach is to add a regularization penalty term to the inverse problem. Regularization is a method for adding constrains, from some a *priori* information or assumption about the structure of f, in addition to those implicit in coherence to the data:

$$\min_f \mathcal{J}(f) \quad subject \quad to \quad \|\mathcal{A}(f) - b\|^2 \leq \sigma^2 \tag{4.2}$$

By formulating the minimization problem using Lagrange multipliers:

$$\widehat{f} = arg\min_f \|\mathcal{A}(f) - b\|^2 + \lambda \mathcal{J}(f) \tag{4.3}$$

Where the first term ensures fidelity to the data, and the second term imposes a roughness regularization penalty $\mathcal{J}$ that is a convex functional of f. The optimal parameter λ is a positive parameter that balances theses two terms and is chosen such that $\left\|\mathcal{A}(\hat{f}) - b\right\|^2 \approx \sigma^2$.

One popular choice for the regularization term includes quadratic penalties [85] that is the squared l_2 norm of either the image f or its (discrete) derivatives:

$$\mathcal{J}(f) = \|\nabla f\|^2 = \int_\Omega |\nabla f|^2 d\Omega \tag{4.4}$$

Above penalty term is well-known as Tikhonov regularization [86, 87]. By using a quadratic regularization term, the estimator $\widehat{f}$ becomes a linear combination of quadratic terms that provides computational advantages. However, quadratic estimators suffer from oversmoothing of recovered image, as they do not recover some important attributes of f, such as the location and magnitude of jumps, or higher order discontinuities [84]. To show this, consider following $1-D$ computation on step edges.

4.2.1.1 1-D Computation on Step Edges

Set $\Omega = [-1, 1]$, and f the step edge function:

$$f(x) = \begin{cases} 0 & \text{x} \leq 0 \\ \text{a} & \text{x} > 0 \end{cases} \tag{4.5}$$

Where a is a real number. Then, the regularization penalty can be written as:

$$\mathcal{J}(f) = \int_{-1}^{1} |f'(x)|^2 \, dx \tag{4.6}$$

Although f is not differentiable at 0, we try to compute the above penalty by approximating $f'(x)$ around 0:

$$\begin{gathered} f'(x) \approx \frac{f(h) - f(-h)}{2h} \approx \frac{a}{2h} \\ s.t. \quad x \in [h, -h] \quad and \quad h > 0, small \end{gathered} \tag{4.7}$$

Then:

$$\begin{aligned} \int_{-1}^{1} |f'(x)|^2 \, dx &= \int_{-1}^{-h} |f'(x)|^2 \, dx + \int_{-h}^{h} |f'(x)|^2 \, dx + \int_{h}^{1} |f'(x)|^2 \, dx \\ &\approx 0 + 2h \times (\frac{a}{2h})^2 + 0 \approx \frac{a^2}{2h} \to \infty, h \to 0 \end{aligned} \tag{4.8}$$

Therefore, a step-edge is severely penalized as it has infinite energy and cannot minimize the Tikhonov regularization. Now replace the square in the regularization term by a $p > 0$:

$$\begin{aligned} &\int_{-1}^{1} |f'(x)|^p \, dx \\ &\approx 0 + 2h \times |\frac{a}{2h}|^p + 0 \approx |a|^p (2h)^{1-p} < \infty, \quad when \quad p \leq 1, \quad h \to 0 \end{aligned} \tag{4.9}$$

Equation (4.9) shows that the regularization term is finite when $p \leq 1$, so edges are less penalized. Note that when $p = 1$, this is the "Total Variation" [88] of

f where the penalty is the l_1 norm of gradient magnitude of signal, i.e.,

$$\mathcal{J}_{TV}(f) = \int_{\Omega} |\nabla f| d\Omega \tag{4.10}$$

The concept of using total variation in image processing was first introduced by Rudin *et al.* [88] for noise removal, since it is very effective at simultaneously preserving edges while smoothing noise in flat regions [89, 90]. In addition, total variation (TV) minimization is widely used for reconstruction of images with sparse gradients that specially makes sense as many natural images have sparse or nearly sparse gradients [91, 92].

Generally, constrained l_1 minimization methods are well-known for reconstruction of sparse signals from highly incomplete sets of linear measurements [93]. This is thoroughly investigated in the next section.

4.2.2 l_1-norm Minimization in Sparse Signal Recovery

To investigate l_1-norm minimization problem, let's take a step out and look at this method in the field of signal recovery. One of challenging problems in engineering is to reconstruct a signal when there are fewer equations than unknowns. Such a problem of course does not have a unique solution without some additional information. However, under sparsity assumption we can find a unique solution which has fewest non-zero entries.

We can recover a signal $x_0 \in \mathbb{R}^n$ by solving the following optimization problem under the sparsity assumption, when the unknown signal that we wish to recover

depends upon a smaller number of unknown parameters:

$$\min_{x\in\mathbb{R}^n} \|\Phi x - y\|^2 + \lambda \|x\|_{l_0} \tag{4.11}$$

Where Φ is an $m \times n$ matrix with fewer rows than columns ($m < n$), and $\|x\|_{l_0} = |\{i : x_i \neq 0\}|$; i.e., number of nonzero samples in x.

l_0 norm is the sparsity count regularization. However, the equation in (4.11) is nonconvex, and a common alternative is to consider the following convex problem:

$$\min_{x\in\mathbb{R}^n} \|\Phi x - y\|^2 + \lambda \|x\|_{l_1} \tag{4.12}$$

Where $\|x\|_{l_1} = \sum_{i=1}^{n} |x_i|$. In equation (4.12), l_1 norm is used as a poxy for the l_0 sparsity count.

4.2.3 Enhance Sparsity Recovery in l_1-norm Minimization

Like l_0 norm, the l_1 norm regularization term has an advantage over quadratic penalty functions as it preserves jumps in the function. However, a key difference between the l_1 and the l_0 norms is the dependence on magnitude, such that larger coefficients are penalized more heavily than the smaller coefficients in the l_1 norm. To address this imbalance, Candes *et al.* have suggested a weighted formulation of l_1 minimization to more democratically penalize nonzero coefficients [93]:

$$\min_{x\in\mathbb{R}^n} \|\Phi x - y\|^2 + \lambda \|Wx\|_{l_1} \tag{4.13}$$

Where $\|Wx\|_{l_1} = \sum_{i=1}^{n} w_i |x_i|$, and each w_i is a positive weight.

In general, weighted and unweighted l_1 minimization have different solutions, since weights can be considered as free parameters in the convex problem, whose

values "can" improve the signal reconstruction if they are set wisely.

What values for the weights will improve signal reconstruction?

As a rough rule of thumb, weights should be chosen such that they counteract the influence of signal magnitude on the l_1 penalty function. Suppose we know the true signal x_0; then, the weights can be inversely proportional to the true signal magnitude such that:

$$w_i = \begin{cases} \frac{1}{|x_{0,i}|}, & x_{0,i} \neq 0, \\ \infty, & x_{0,i} = 0. \end{cases} \tag{4.14}$$

Ideally, zero-valued components of x_0 are prohibited in the recovered signal being penalized by large (infinite) entries of w_i, while the largest signal coefficients are encouraged to be identified as nonzero in the recovered signal being penalized less by small (finite) w_i entries. It is of course impossible to construct the precise weights without knowing the true signal x_0, so a valid set of weights can be designed based on an approximation $\hat{x_0}$ to x_0. Also, To provide stability, the equation (4.14) is rewritten as following by introducing the parameter $\epsilon > 0$:

$$w_i = \frac{1}{|\hat{x_{0,i}}| + \epsilon}, \quad \epsilon > 0 \tag{4.15}$$

Where ϵ should be set slightly smaller than the expected nonzero magnitudes of $\hat{x_0}$ [93]. In addition to stability, using ϵ ensures that a zero-valued entry in approximated $\hat{x_0}$ does not strictly prohibit a nonzero estimate in the recovered signal.

Using above weights in minimization problem (4.13) makes the solution x to concentrate on the large nonzero entries in x_0 since they will be penalized less by small w_i entries, while zero-valued entries on x_0 will be largely penalized and are

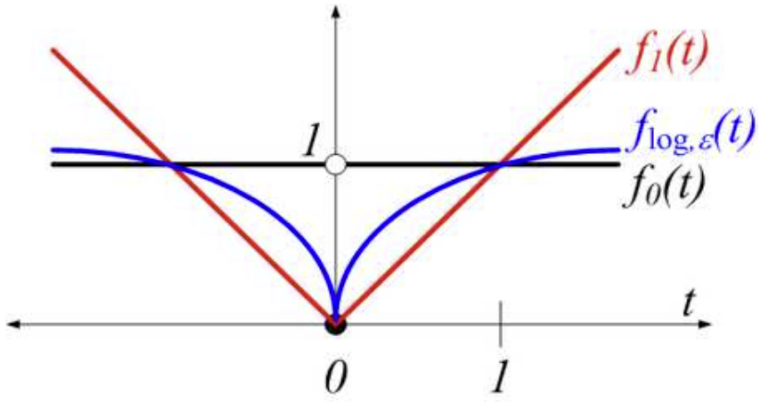


Figure 4.1: The log-sum concave penalty function $f_{log,\epsilon}(t)$ is a better approximation for the l_0 sparsity count $f_0(t)$ rather than the traditional convex l_1 regularization $f_1(t)$ (Candes *et al.*, 2008).

discouraged in the recovered signal.

4.2.4 Analytical Justification

Using a concave penalty function, instead of l_1-norm regularization term in equation (4.12), more closely resembles the l_0-norm regularization. It is illustrated in figure (4.1), where the $f_{log,\epsilon}(t)$ is defined as:

$$f_{log,\epsilon}(t) = log(1 + \frac{|t|}{\epsilon}) \tag{4.16}$$

Like the l_0 norm, the $f_{log,\epsilon}(t)$ allows a relatively large penalty to be placed on small nonzero coefficients. In fact, $f_{log,\epsilon}(t)$ tends to $f_0(t)$ as $\epsilon \to 0$.

In this section, we show that using a weighted l_1 minimization of as weights defined in (4.15) is like to find the local minimum of a concave penalty function as

defined in (4.16).

To establish this connection, consider the following problem:

$$\min_{x\in\mathbb{R}^n} \|\Phi x - y\|^2 + \lambda \sum_{i=1}^{n} log(1 + \frac{|x_i|}{\epsilon}) \tag{4.17}$$

Which is equivalent to

$$\min_{x\in\mathbb{R}^n} \sum_{i=1}^{n} log(1 + \frac{|x_i|}{\epsilon}) \quad subject \quad to \quad y = \Phi x \tag{4.18}$$

Above optimization problem can be solved using a majorize-minimize (MM) [94] framework by iteratively minimizing a simple surrogate function majorizing a given objective function, so (4.18) is equivalent to:

$$\min_{x,u\in\mathbb{R}^n} \sum_{i=1}^{n} log(1 + \frac{u_i}{\epsilon}) \quad subject \quad to \quad \begin{array}{c} y = \Phi x, \\ |x_i| \leq u_i, i = 1, .., n. \end{array} \tag{4.19}$$

If $\hat{x}$ is a solution to (4.18), then $(\hat{x}, |\hat{x}|)$ is a solution to (4.19). Also, conversely, if $(\hat{x}, \hat{u})$ is a solution to (4.19), then $\hat{x}$ is a solution to (4.18). Now, set:

$$g\left(u\right) = \sum_{i=1}^{n} log(1 + \frac{u_i}{\epsilon}) \tag{4.20}$$

Function g is concave and below its tangent, so it can be minimized by iteratively improving on an initial guess $u^{(0)}$. At each iteration, we minimize a linear approximation of g around the previous guess $u^{(l)}$ derived from the first-order Taylor polynomial:

$$u^{(l+1)} = arg\min \left\{ g\left(u^{(l)}\right) + \bigtriangledown g\left(u^{(l)}\right) . \left(u - u^{(l)}\right) \right\} \quad subject \quad to \quad u \in \mathcal{C} \tag{4.21}$$

Where $\mathcal{C}$ is a convex set. Each iteration of above problem is a convex optimization problem, since it is minimization of a linearization of g around previous

guess:

$$u^{(l+1)} = arg\min \bigtriangledown g\left(u^{(l)}\right).u \quad subject \quad to \quad u \in \mathcal{C} \tag{4.22}$$

In the case of optimization problem in (4.19), this gives:

$$\left(x^{(l+1)}, u^{(l+1)}\right) = arg\min \sum_{i=1}^{n} \frac{u_i}{u_i^{(l)} + \epsilon} \quad subject \quad to \quad \begin{matrix} y = \Phi x, \\ |x_i| \leq u_i, i = 1, .., n. \end{matrix} \tag{4.23}$$

Which is equivalent to:

$$x^{(l+1)} = arg\min \sum_{i=1}^{n} \frac{|x_i|}{\left|x_i^{(l)}\right| + \epsilon} \quad subject \quad to \quad y = \Phi x \tag{4.24}$$

By setting:

$$W_i^{(l+1)} = \frac{1}{\left|x_i^{(l)}\right| + \epsilon} \tag{4.25}$$

Then (4.24) can be rewritten as:

$$x^{(l+1)} = arg\min \sum_{i=1}^{n} W_i^{(l+1)} |x_i| \quad subject \quad to \quad y = \Phi x \tag{4.26}$$

That is equivalent to:

$$x^{(l+1)} = arg\min \|\Phi x - y\|^2 + \lambda \left\|W^{(l+1)} x\right\|_{l_1} \tag{4.27}$$

Above is an iterative reweighted l_1 minimization approach suggested by Candes *et al.* [93], in which each iteration of algorithm solves a convex optimization problem, whereas the overall algorithm finds a local minimum of a concave penalty function.

If we suppose weights are pre-specified by a "true knowledge" of signal x_0; then, only one iteration of above algorithm would be enough:

$$\begin{gathered} \hat{x_0} = arg\min \|\Phi x - y\|^2 + \lambda \|Wx\|_{l_1}, \\ where \quad W = \frac{1}{|x_0| + \epsilon} \end{gathered} \tag{4.28}$$

4.2.5 Variation of Weights

Non-convex regularization terms give reconstruction with less blurring than convex metrics [95]. As shown in section 4.2.4, using pre-specified weights in a weighted l_1 minimization is like to find the local minimum of a concave penalty function that is a better approximation of l_0 norm (see Figure (4.1)).

Depending on selected concave function, there are a variety of possible weighting functions in place of W as defined in (4.28). For example, if instead of a log-sum penalty function, as defined in (4.20), we consider an arctangent concave function:

$$g(u) = \sum_{i=1}^{n} atan(\frac{u_i}{\epsilon}) \tag{4.29}$$

We can find an alternative formulation of spatial weights by following the same procedure described in section 4.2.4:

$$W = \frac{1}{x_0^2 + \epsilon^2} \tag{4.30}$$

The choice of different variations of weighting function can be the subject of further empirical studies. The results of this study are provided based on the choice of weighting function as defined in (4.28).

4.3 Methods

4.3.1 Weighted Total Variation Minimization for Image Reconstruction

To achieve equation (4.28), we supposed that we have a true knowledge of signal x_0. Surprisingly it can be a valid assumption in medical imaging since we may have different representations of current subject image through different modal-

ity sources, where the other modality scans may provide a better estimate of some anatomical features that we wish to recover in the current subject image.

The weighted l_1-norm minimization concept, introduced in section 4.2.3, can also enhance the performance of total-variation (TV) minimization in image reconstruction, since TV can be considered as an l_1-norm minimization problem. Weighted-TV problem can then be presented as:

$$\widehat{f} = arg\min_{f} \|\mathcal{A}(f) - b\|^2 + \lambda \|W\,|\bigtriangledown f|\|_{l_1}\,, \qquad where \quad W = \frac{1}{\left|\bigtriangledown \widehat{f}_0\right| + \epsilon} \tag{4.31}$$

Where $\mathcal{A}$ is a Fourier undersampling operator modeling the physical process that causes degradation; b is the obeservation samples presented as a vector of noisy low-pass Fourier measurements, and $|\bigtriangledown f|$ denotes the magnitude of discrete gradients of f.

Also, $\left|\bigtriangledown \widehat{f}_0\right|$ is an estimation of gradient magnitude of f computed as a prior edge map from the high-resolution representation of input image subject in other modality scans. The spatial weights (W) are then constructed from this estimated edge map.

Weights are inversely proportional to the estimation of gradient magnitude of input image, such that the strongest edges get the lowest weight values close to zero, and weak edges get higher weights. The maximum weight value is assigned to the smooth regions with no edges.

Some methods have been developed for accurate estimation of spatial weights. Candes *et al.* [93] suggested an iterative framework to estimate the weights iteratively

from the gradient magnitude of input low-resolution image. First, all weights are set to one; then, at each iteration weights are updated based on the gradient magnitude of estimated high-resolution image in previous iteration. Ongie and Jacob (2015) [96] suggested to expand the theory of sampling signals of finite rate of innovation (FRI) on the input low-resolution image to estimate a resolution-independent mask whose zeros represent the edges of image; then, the spatial weights are created by discretizing the estimated mask at desired resolution.

Here, we create the spatial weights based on a high-resolution edge map estimated from equivalent representation of underlying anatomical structures in other modality sources. Following section describes the suggested multi-modal framework to estimate a combined edge map from different modality scans provided in higher spatial resolution. Then, spatial weights are constructed to be inversely proportional to the values of estimated edge map.

Several optimization methods have been developed [97, 98, 99] to solve the minimization problem in equation (4.31) efficiently.

4.3.2 Construction of Weights from Estimated Anatomical Edge Map

Here we describe a multi-modal framework to construct the spatial weights from an estimation of the anatomical boundaries through a combined edge map. The edge map is derived from other modality scans that are better representation of the underlying anatomical structure in higher spatial resolution. Then, the spatial weights are created to be inversely proportional to the edge values.

Assume $I_i, i \in \{1, ..., N\}$ are high-resolution modality scans all representing a single subject image. The gradient of each image is defined as:

$$\forall x = (x_1, ..., x_n) \in \{voxel \quad locations\}\,,$$
$$g_i(x) = \nabla I_i(x) = \begin{bmatrix} \frac{\partial I_i}{\partial x_1}(x_1, ..., x_n) \\ . \\ . \\ . \\ \frac{\partial I_i}{\partial x_n}(x_1, ..., x_n) \end{bmatrix} \tag{4.32}$$

Where $g_i(x)$ is the gradient of i^{th} input scan at voxel location x. Consequently, the gradient magnitude of each image is defined as:

$$|g_i(x)| = \|\nabla I_i(x)\|_2 = \sqrt{(\frac{\partial I_i}{\partial x_1}(x_1, ..., x_n))^2 + ... + (\frac{\partial I_i}{\partial x_n}(x_1, ..., x_n))^2} \tag{4.33}$$

Then, the edge map of underlying anatomical structure is inferred from gradient magnitudes of all input multi-modal scans:

$$\forall x \in \{voxel \quad locations\}\,,$$
$$\mu(x) = \max_x \{T_i\,[g_i(x)]\} \tag{4.34}$$

Where T_i is an image intensity transformation function defined as:

$$T_i(I) = \begin{cases} \alpha_i I + b_i & \text{if } Q_I(50) < I < Q_I(95) \\ M & \text{if } I > Q_I(95) \\ \epsilon & \text{if } I < Q_I(50) \end{cases} \tag{4.35}$$
$$where \quad \alpha_i = \frac{M - \epsilon}{Q_I(95) - Q_I(50)}\,,\; b_i = M - \alpha_i . Q_I(95)$$

Where $Q_I(p)$ is the p^{th} quantile of input intensity range. M is maximum mapped value. Here we set M to 255 that is the maximum allowed value for unsigned

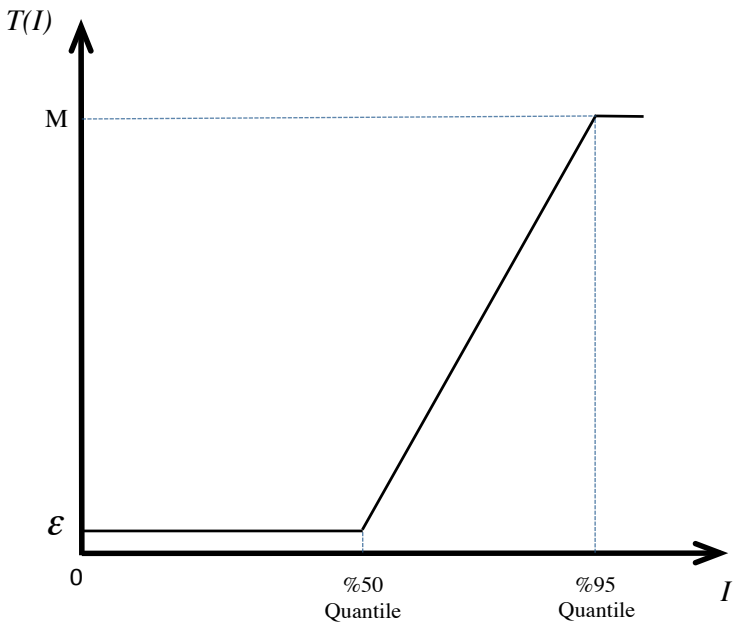


Figure 4.2: Intensity transform function. Strong edges (with values above %95 percentile) are mapped to a maximum value M. Weak edges (with values below %50 percentile) are mapped to a value $\epsilon > 0$ close to zero. Other edges are mapped linearly to a $\epsilon < value < M$.

short. Finally, $\epsilon > 0$ is the minimum mapped value. Here, ϵ is set to 1, that is the minimum non-zero value for the unsigned short range.

Figure (4.2) shows the intensity transformation function defined above. In fact, the designed transform function maps the strongest edges to have the same maximum value and removes all weak edges below the median intensity value. Then, at each voxel location, the edge value is selected from the modality scan that provides the maximum contrast in that location (equation (4.34)).

Table 4.1: Test dataset for initial assessments.

Scan	Site	MR vendor	Field strength	Collected modalities
1166_54860	Site_001 (Rochester)	GE Signa HDxt	3.0	T1: $1 \times 1 \times 1\ mm^3$ T2: $1 \times 1 \times 1\ mm^3$ DWI: $1 \times 1 \times 2.4\ mm^3$

Once the edge map μ is estimated, the spatial weights are defined as:

$$W(x) = \frac{1}{\mu(x)} \tag{4.36}$$

4.4 Preliminary Assessments on 2D Data

This section presents a series of experiments on a two-dimensional (2D) image data for the initial assessments of developed super-resolution reconstruction approach. The results compared the performance of our SRR method to standard TV and zero-padded IFFT approaches.

4.4.1 Test Data

Initial assessments were run on two-dimensional data. For this purpose, a sample 3D dataset, listed in Table (4.1), was selected to create the input 2D test images. This dataset was used as part of multi-site international PREDICT-HD project and was acquired with a high isotropic ($1 \times 1\ mm^2$) resolution in the axial plane.

The input baseline and 2D test images were then created as follows. Note that all inputs are aligned in both physical and voxel spaces.

Baseline 2D DWI image: First, the $b0$ component and the 1^{st} gradient component were extracted from the sample 4D diffusion-weighted dataset listed in table (4.1). Then, the mid-axial slice of these components were extracted to serve as our high-resolution ground truth images with the size of 256×256 and isotropic pixel sizes of $1 \times 1\ mm^2$.

Low resolution input DWIs: were created by downsampling the high-resolution baseline images by a factor of 2 by only keeping the low pass Fourier indices from the high-resolution image.

Structural MRI input images: were created by extracting the mid-axial slice from the corresponding $T1/T2$-weighted images of same data session.

4.4.2 Evaluation Metric

As a general approach similar to other super-resolution studies [76, 81, 96], signal-to-noise ratio (SNR) was used to compare the output of developed SRR method to the results from standard TV and zero-padded IFFT approaches, such that the *higher* SNR value indicates the better reconstruction performance. For each reconstructed image from the different methods, the SNR was computed as:

$$SNR = 20 \times \log_{10} \frac{\|I_0\|_2}{\|I_{SR} - I_0\|_2} \tag{4.37}$$

Where I_{SR} is the reconstructed super-resolved DWI image, and I_0 is the baseline high-resolution DWI image.

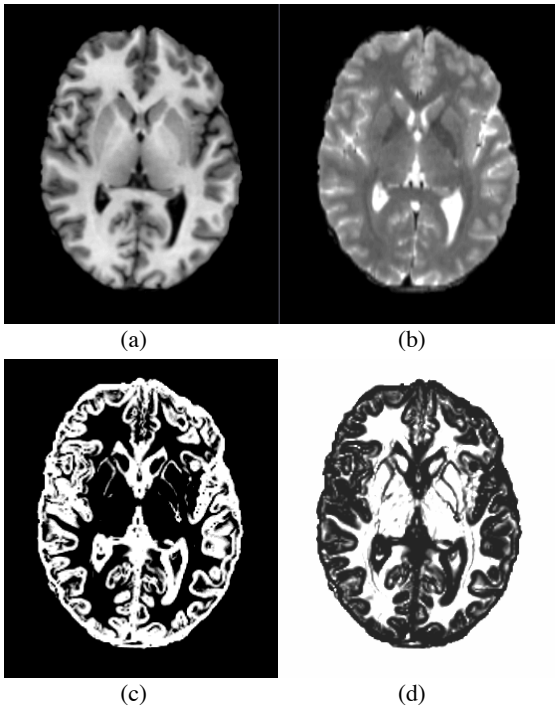


Figure 4.3: Created spatial weights based on the anatomical edges estimated from the high-resolution structural MR modalities. (a) T1-weighted MR image. (b) T2-weighted MR image. (c) Estimated edge map based on the maximum gradient values as described in section 4.3.2. (d) Prior spatial weights created from estimated edge map based on equation (4.36).

4.4.3 Preliminary 2D Results

Figure (4.3) shows the $T1/T2$-weighted images, the estimated anatomical edge map, and the created spatial weights based on the method described in section 4.3.2.

The prior spatial weights map was passed to a weighted-TV algorithm along with the input low-resolution 2D test images to create the results of developed method. Developed approach was tested on both $b0$ and 1^{st} gradient components from a

4D DWI dataset. Figure (4.4) shows the performance of our weighted-TV super-resolution reconstruction method compared to standard TV and zero-padded IFFT on a low-resolution $b0$ image. Figure (4.5) shows the same results for the 1^{st} gradient component image. We also provided the reconstruction results using FRI edge map suggest by Jacob and Ongie [96] demonstrated in Figure (4.6). The FRI method performance is also superior to zero-padded IFFT and standard TV algorithms, and its reconstructed output image is comparable to the results of our developed method.

4.4.4 Conclusion

The developed weighted-TV method based on the prior knowledge of anatomical edges shows superior performance compared to both standard TV and zero-padded IFFT for both b_0 and first gradient component.

Also, zero-padded IFFT approach gets better results rather than standard TV for the b_0 component. The reason is that TV works well for compressed sensing style sampling (when we have sparse samples equally from low and high frequencies) but seems to perform poorly for super-resolution (when mostly higher frequencies are missed). This is specially the case here, since our input real data is a relatively smooth image with no super sharp edges and less high frequency content.

The FRI algorithm has an advantage over our developed method, since it estimates the anatomical edges from the input low-resolution image, and unlike our method, it does not need complementary information from external sources. However, the developed approach has following advantages:

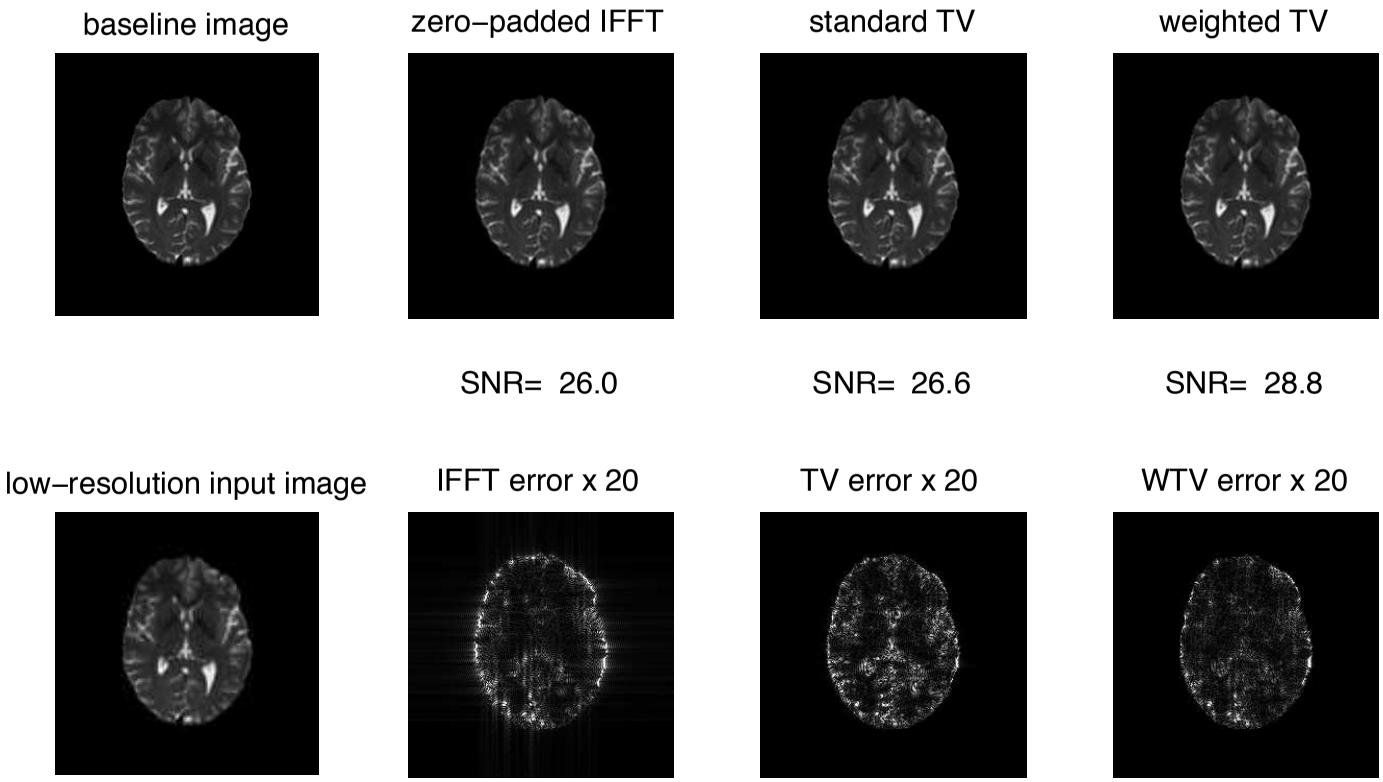


Figure 4.4: Reconstructed b_0 component of a typical DWI subject using 3 different methods. The SNR value is provided for the result of each method demonstrated in the first row. The second row shows the low-resolution image, by a factor of 2, and the difference image between each reconstructed image and the original high-resolution baseline image, that is magnified by 20 times. The higher SNR value for the developed weighted-TV approach indicates better reconstruction results.

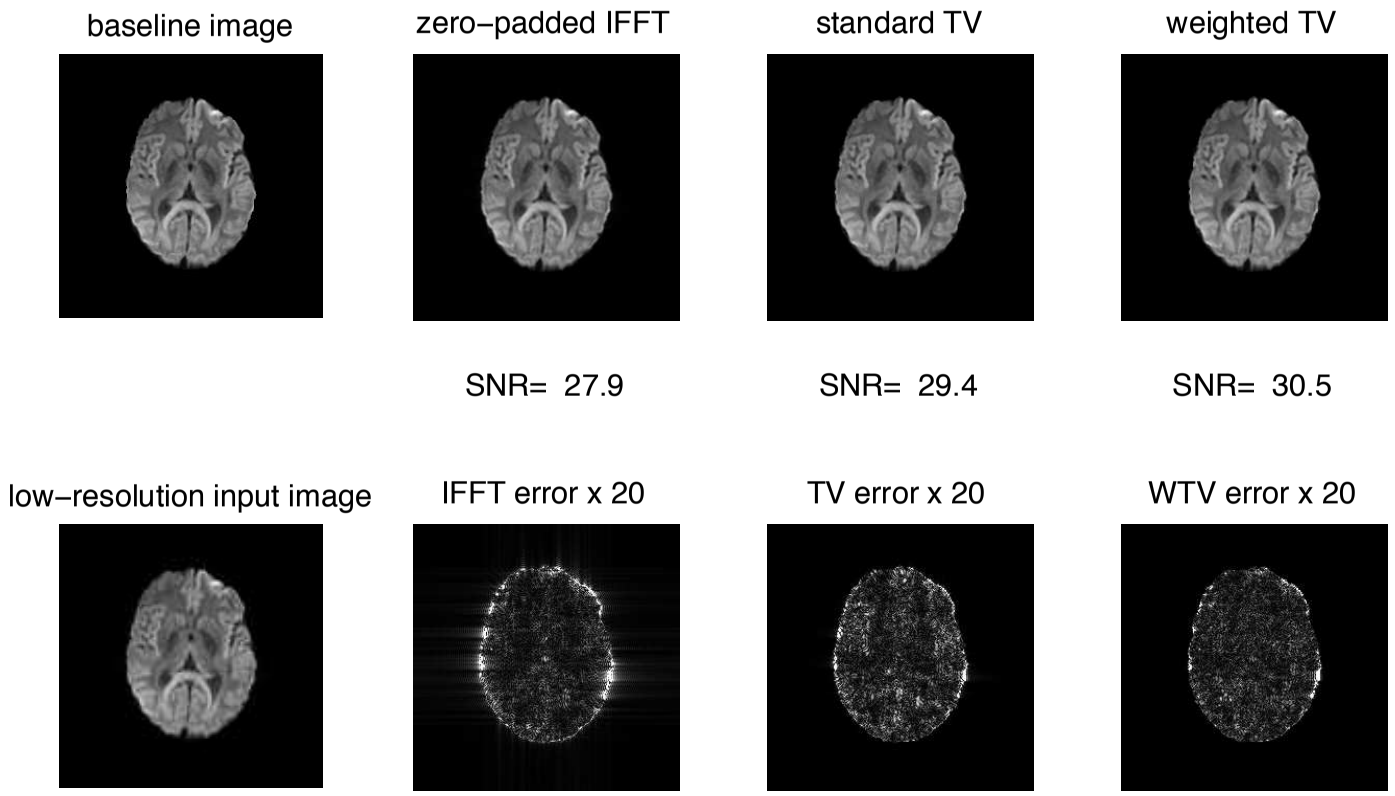


Figure 4.5: Reconstructed first gradient component of a typical DWI subject using 3 different methods. The SNR value is provided for the result of each method demonstrated in the first row. The second row shows the low-resolution image, by a factor of 2, and the difference image between each reconstructed image and the original high-resolution baseline image, that is magnified by 20 times. The higher SNR value for the developed weighted-TV approach indicates better reconstruction results for the gradient components as well.

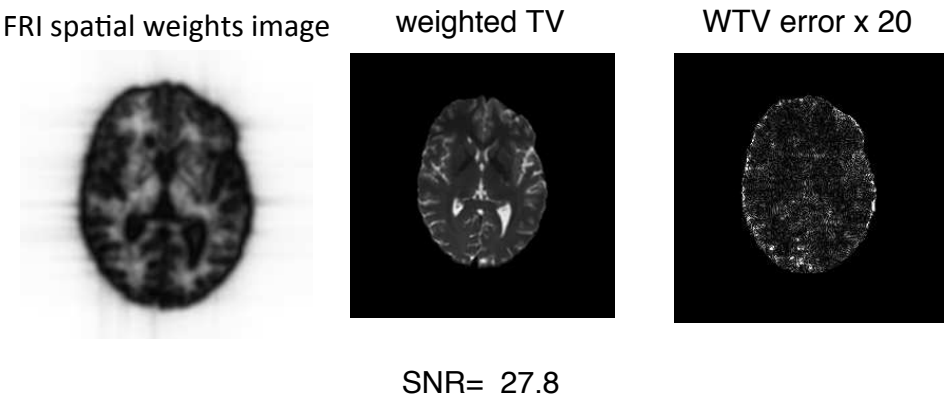


Figure 4.6: Reconstructed b0 component of a typical DWI subject using FRI edge map suggest in (Ongie and Jacob, 2015).

- Using conventional image processing filters makes our method to be easily adoptable for the processing of large-scale multi-site datasets, since there are less parameters needed to be adjusted to work optimally for the analysis of a set of heterogeneous data.
- Developed method uses fast conventional image processing tools. It accelerates the process time that can be especially important when the implementation is expanded to operate on real 3D datasets.
- The introduce FRI algorithm operates in voxel space; however, the developed method can be easily expanded to operate in physical space, that is important in real world medical imaging applications.

4.5 Experimental Methods and Results

The developed novel edge-guided weighted total variation (WTV) super-resolution reconstruction (SRR) method operates on 3D images that have accurate physical space representations. This 3-dimensional method is incorporated into a framework that allows processing of all 3D sub-volumes of a 4D DWI dataset in a fully automated processing pipeline.

The developed novel SRR method is evaluated using *three* different quantitative evaluation approaches and is compared with two other methods in the literature: zero-padded IFFT, and standard total variation (TV).

4.5.1 Test Data

We identified 20 DWI sessions from WU-Minn Human Connectome Project (HCP) database [100, 101] that can provide ground truth high-resolution (HR) baselines.

The HCP designed $3T$ Siemens Connectome scanner equipped with 100 mTm^{-1} and 300 mTm^{-1} gradient coils, that are several times more powerful than standard clinical scanners, and exploited several imaging and image reconstruction innovations to speed up acquisition and improve the data quality [101].

The HCP diffusion-weighted imaging data were acquired with a resolution of $1.25 \times 1.25 \times 1.25$ mm^3 with 18 $b0$ components and 90 gradient directions at $b = 2000$ s/mm^2. The original HCP data were used as high-resolution baselines. We generated low-resolution datasets from these HR baselines to be used as inputs

for comparing different super-resolution recovery techniques. To simulate a group of typical-resolution DWI as input sources, the original HCP data were downsampled in Fourier domain by a factor of 2 to obtain DWI with resolution of $2.5 \times 2.5 \times 2.5$ mm^3 that is at the similar level of our typical DWI resolution.

The quantitative evaluation approaches are listed as follows:

4.5.2 **Approach 1**: Evaluation Based on the Overlap Between Tractography Results

We evaluated the performance of the developed SRR method in providing an accurate assessment of brain connectivity by extracting the cortico-spinal tract (CST) and arcuate fascicle (AF) in both HR baseline and reconstructed images and computing the fiber bundle overlap between the extracted tracts from the HR baseline and each reconstructed image.

Arcuate fascicle (AF) is located ventral to the superior longitudinal fascicle II (SLF II) [102] and dorsal to the extreme capsule and the superior circular sulcus of the insula. It connects the inferior frontal gyrus with the middle temporal gyrus, the posterior part of the superior temporal gyrus and cortices of the lateral temporo-occipital transition region [103, 104]. The cortico-spinal tract (CST) consists of fibers that originate in the paracentral lobule (principally the precentral gyrus), course through the brainstem, pons and medullary pyramid, descending into the spinal cord. CST constitutes the most prominent descending fiber system of the central nervous system [105, 102]. Figure (4.7) shows the tractgraphy results for AF from right view

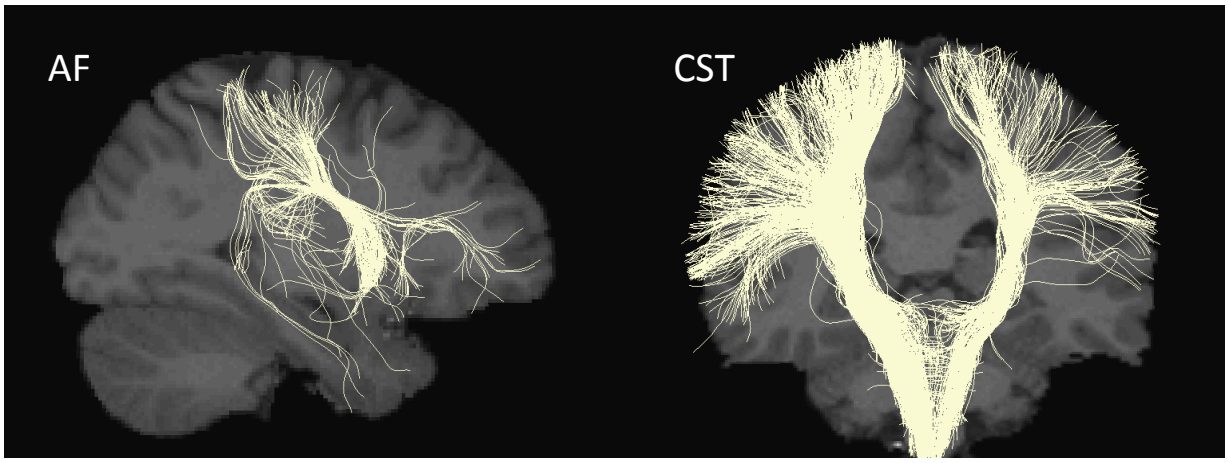


Figure 4.7: (Left): Arcuate Fascicle (AF) from right view. (Right): Cortico-spinal tract (CST) from anterior view.

and for CST from anterior view.

First, a whole brain multi-tensor tractography [106, 107] was computed on the HR gold-standard scan and each reconstructed image. Then, the cortico-spinal tract (CST) and arcuate fascicle (AF) were extracted using the White Matter Query Language (WMQL) [102], that is a technique to formally describe white matter tracts and to automatically extract them using Freesurfer cortical parcellations [108, 109]. This query language allows constructing an anatomical definition for each white matter tract including description of adjacent gray and white matter regions and rules for the spatial relations. Therefore, tracts of interest can be extracted from anatomical knowledge of human brain white matter. Finally, the Bhattacharyya coefficient [110] was used to quantify the fiber bundle overlap. This metric ranges from 0 to 1, with 0 being no overlap, and 1 being complete overlap of fibers.

Bhattacharyya coefficient (BC) is computed based on the probability distributions of each of the spatial coordinates (x, y, z) of a fiber bundle [111]. Denote BC_x as the Bhattacharyya metric on the x-coordinate:

$$BC_x = \int \sqrt{p_b(x)p_{sr}(x)}dx \tag{4.38}$$

Where $p_b(x)$ is the probability distribution function (pdf) on the x coordinate of the HR baseline image, and $p_{sr}(x)$ is computed from the fiber bundle of SR reconstructed image. Then, we compute the overlap between two fiber bundles by taking the average of all three coordinates:

$$BC = \frac{1}{3}(BC_x + BC_y + BC_z) \tag{4.39}$$

The values of BC are bounded between 0 and 1 that indicates no overlap for 0 and a perfect match for 1. Bhattacharyya coefficient has the advantage of being sensitive to minor deviation in tracts since the probability distributions are smooth [111]. Figure 4.8 shows a sample computation of Bhattacharyya coefficient based on the probability distribution function (pdf) of HR baseline image and SR reconstructed image on each of x, y and z coordinates.

Figure 4.9 shows the Bhattacharryya coefficients computed to quantify the overlap between the tractography results extracted from the HR baseline and each of reconstructed images by the developed WTV, standard TV, and zero-padded IFFT methods. Results are presented for CST and AF tract bundles in each of 20 test subject. Numeric results are summarized in Table 4.2, where we report the average Bhattacharryya coefficients over all test subjects for each tract of interest. The results

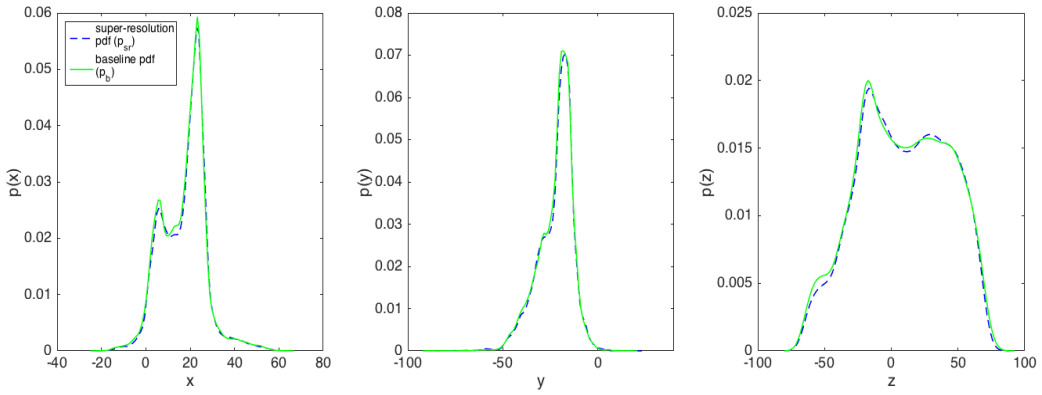


Figure 4.8: A sample computation of Bhattacharyya coefficient based on the probability distribution function (pdf) on each of x, y and z coordinates. In this example BC is calculated as 0.98.

demonstrate that the developed WTV approach shows significant improvement on the tractography results over the other two approaches (standard TV and zero-padded IFFT) for both arcuate fascicle (AF) and cortico-spinal tract (CST). In order to compare the quantitative results, a paired t-test was performed and $p - value < 0.05$ was considered significant. A Welch's t-test or unequal variance t-test was used as it is more reliable than the Student's t-test when the two samples have unequal variances [112]. A two-tailed test was performed to test the null hypothesis that the two population means are equal.

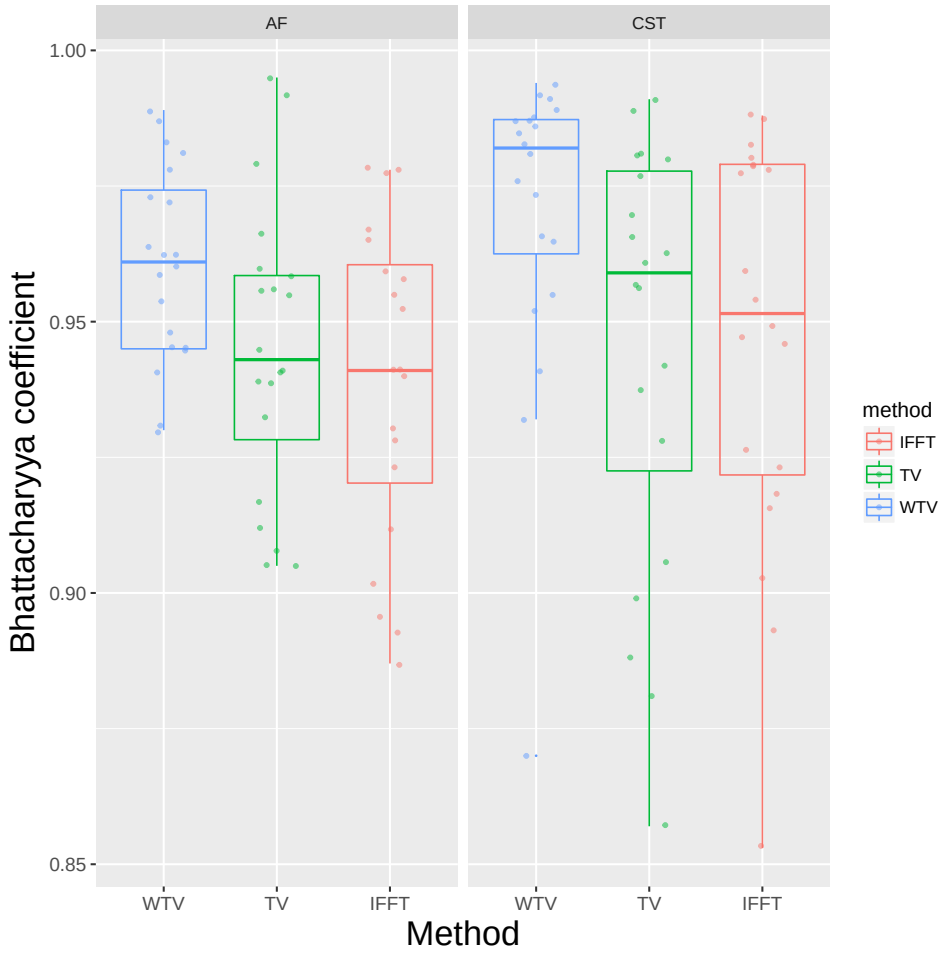


Figure 4.9: The Bhattacharryya coefficients computed for 20 test datasets to quantify the overlap for cortico-spinal tract (CST) and arcuate fascicle (AF) extracted from the HR baseline and each of reconstructed images by three different methods: The developed WTV, Standard TV, and zero-padded IFFT.

Table 4.2: The average Bhattacharyya Coefficient to quantify overlap between tracts of interest extracted from the high-resolution baseline image and each of reconstructed images.

Tract	Method	Average Bhattacharyya Coefficient	p-value vs. WTV
CST	WTV	0.970	
	TV	0.946	$*$
	IFFT	0.947	$*$
AF	WTV	0.960	
	TV	0.945	$*$
	IFFT	0.939	$**$

NOTE: CST: cortico-spinal tracts. AF: arcuate fasciculus tracts. $*$: p-value < 0.05. $**$: p-value < 0.01. (p-value < 0.05 is considered significant).

4.5.3 **Approach 2**: Evaluation Based on the Difference Between Rotationally Invariant Scalar Measurements

We evaluated the performance of the developed SRR method in recovering the high-resolution rotationally invariant scalar (RIS) measurements by computing the average difference of RIS values between the HR baseline and the reconstructed images within four white matter regions of interest. The white matter regions of interest are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well. Pure and non-pure samples are computed from a pure plugs mask that is generated from the low-resolution average $b0$ image and the structural $T1$ and $T2$-weighted MR scans as described in chapter 3.

Figures 4.10, 4.11, 4.12, and 4.13 present the box plots of average RIS errors within each region of interest for each of super-resolution reconstruction (SRR) meth-

Table 4.3: Error Mean Value of DTI scalars within each region of interest across all subjects.

RIS	Region	WTV Mean Error	TV Mean Error	IFFT Mean Error	p-value WTV vs. TV	p-value WTV vs. IFFT
FA	Frontal	2.83×10^{-2}	2.96×10^{-2}	2.96×10^{-2}	$***$	$***$
	Occipital	3.95×10^{-2}	3.97×10^{-2}	3.96×10^{-2}		
	Parietal	3.13×10^{-2}	3.27×10^{-2}	3.26×10^{-2}	$**$	$**$
	Temporal	3.19×10^{-2}	3.37×10^{-2}	3.36×10^{-2}	$***$	$**$
MD	Frontal	2.61×10^{-5}	3.06×10^{-5}	3.07×10^{-5}	$***$	$***$
	Occipital	2.75×10^{-5}	3.24×10^{-5}	3.25×10^{-5}	$***$	$***$
	Parietal	2.82×10^{-5}	3.40×10^{-5}	3.41×10^{-5}	$***$	$***$
	Temporal	2.60×10^{-5}	2.69×10^{-5}	2.69×10^{-5}		
RD	Frontal	2.63×10^{-5}	3.01×10^{-5}	3.02×10^{-5}	$***$	$***$
	Occipital	2.90×10^{-5}	3.28×10^{-5}	3.28×10^{-5}	$***$	$***$
	Parietal	$2..74\times10^{-5}$	3.23×10^{-5}	3.23×10^{-5}	$***$	$***$
	Temporal	2.62×10^{-5}	2.76×10^{-5}	2.76×10^{-5}	$**$	$**$
AD	Frontal	4.11×10^{-5}	4.52×10^{-5}	4.53×10^{-5}	$***$	$***$
	Occipital	5.34×10^{-5}	5.72×10^{-5}	5.71×10^{-5}	$*$	$*$
	Parietal	4.95×10^{-5}	5.47×10^{-5}	5.48×10^{-5}	$***$	$***$
	Temporal	4.52×10^{-5}	4.55×10^{-5}	4.55×10^{-5}		

NOTE: $*$: p-value < 0.05, $**$: p-value < 0.01, $***$: p-value < 0.001, $****$: p-value < 0.0001.

ods. Mean errors are computed for each of rotationally invariant scalars consisting of fractional anisotropy, mean diffusivity, radial diffusivity, and axial diffusivity (FA, MD, RD, and AD) that are defined in section 1.4.3. Error metric in each region of interest is defined as the average difference between the RIS values over all voxels within that region computed from the HR baseline and each of reconstructed images from different SRR methods. All numeric results are summarized in table 4.3, where error metrics are averaged over all test subjects.

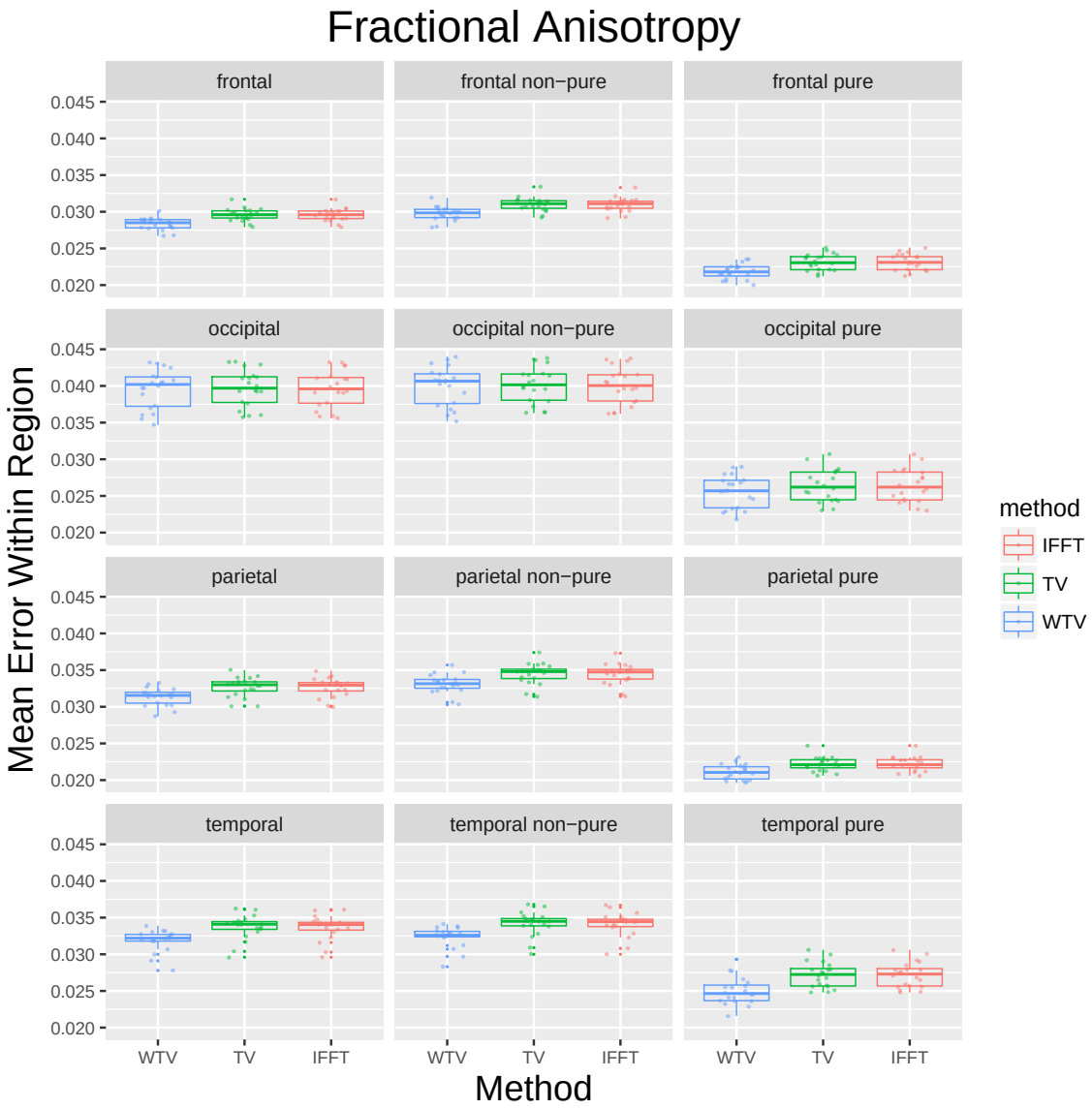


Figure 4.10: Comparing the average fractional anisotropy (FA) values within 4 WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT (IFFT). The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well.

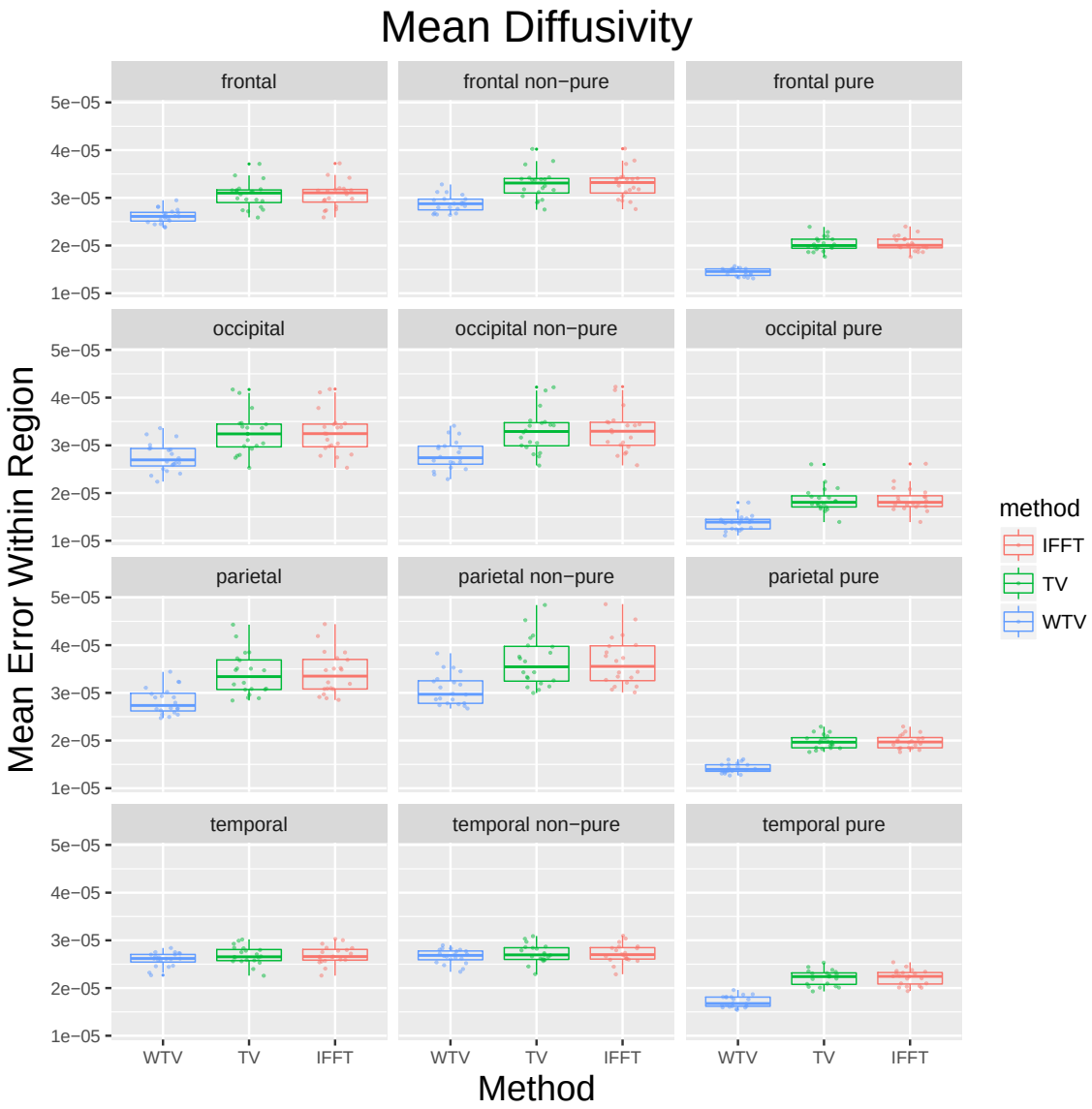


Figure 4.11: Comparing the average mean diffusivity (MD) values within 4 WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT (IFFT). The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well.

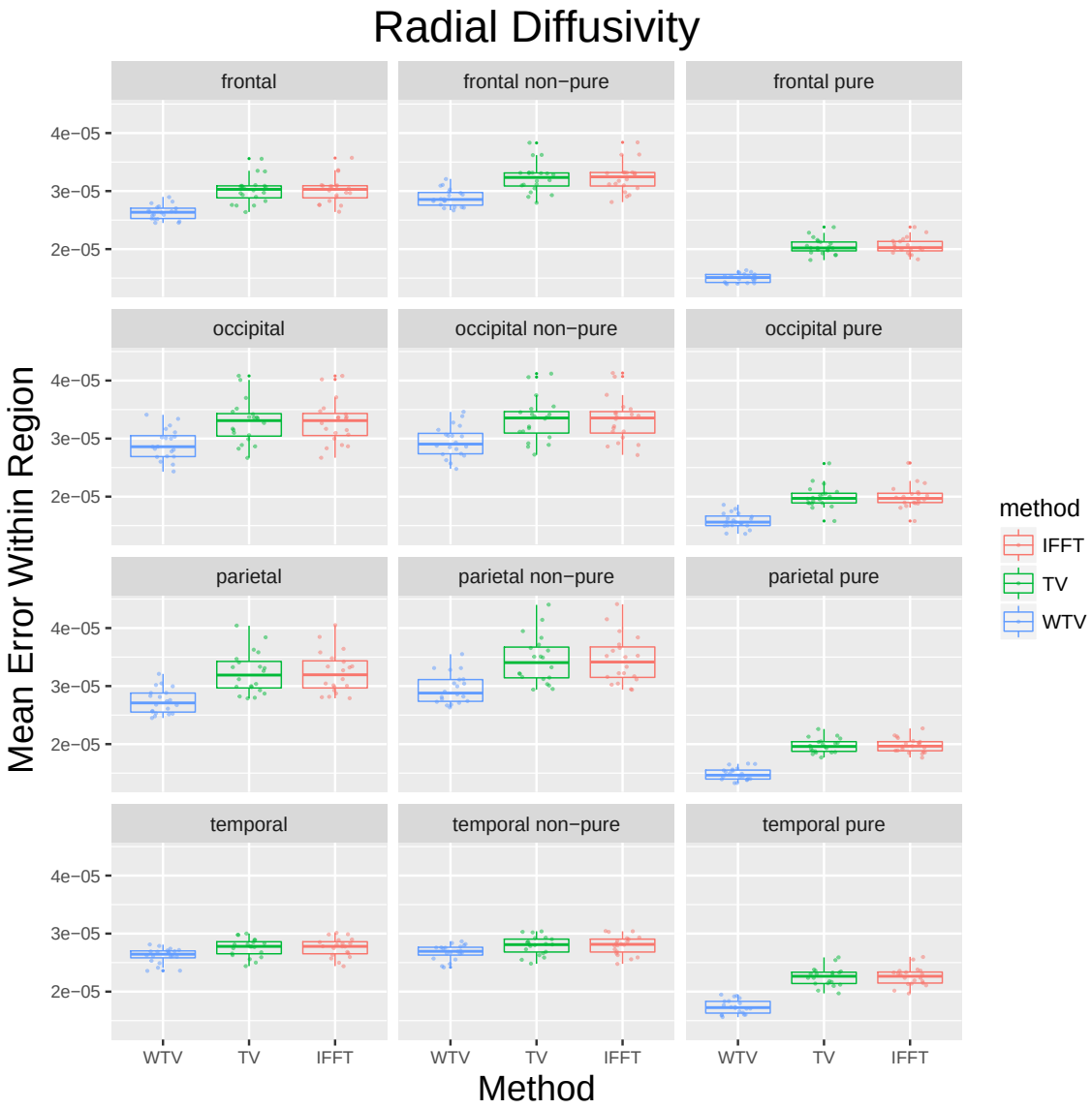


Figure 4.12: Comparing the average radial diffusivity (RD) values within four WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT (IFFT). The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well.

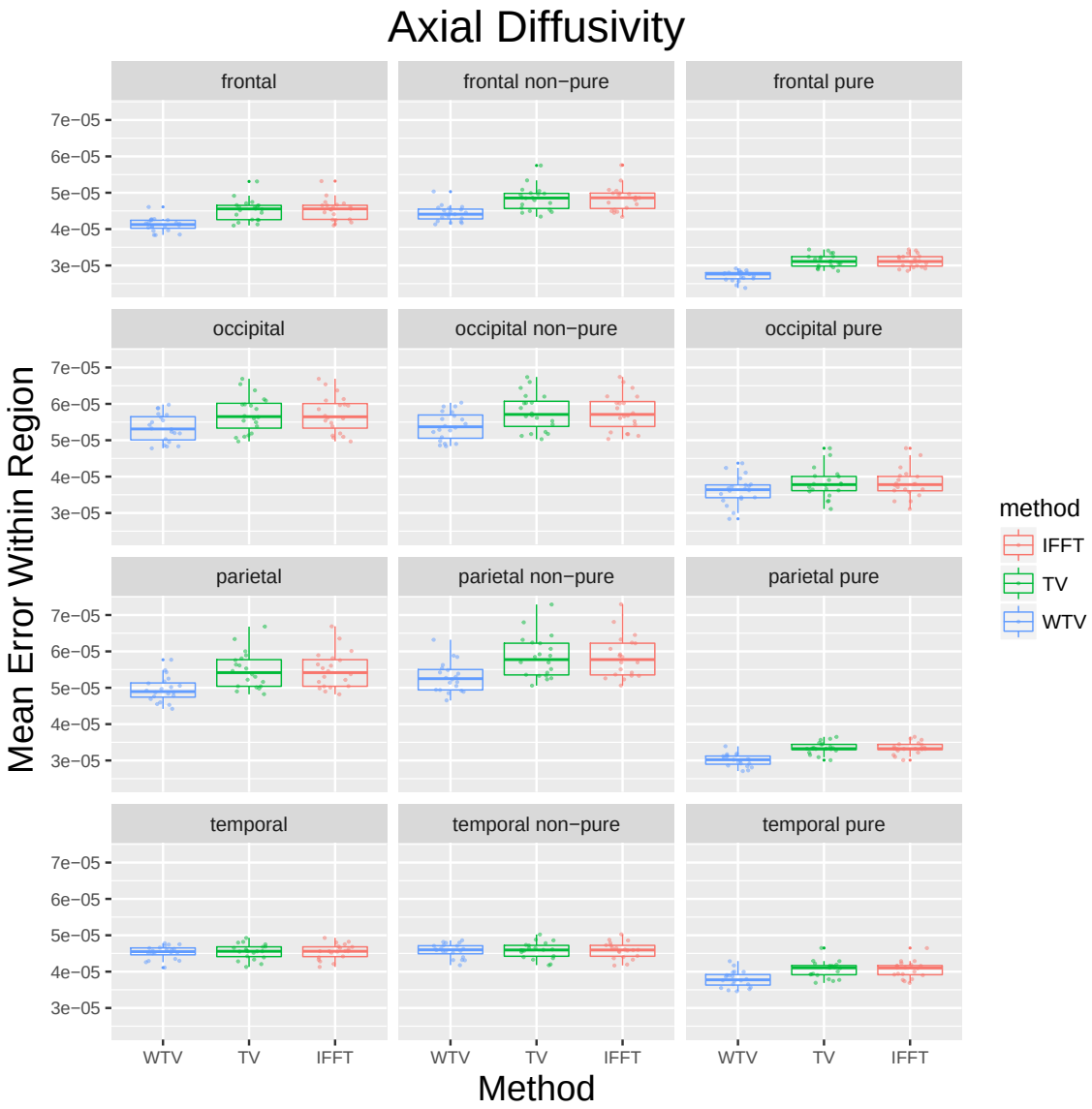


Figure 4.13: Comparing the average of axial diffusivity (AD) values within 4 WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT (IFFT). The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well.

Standard TV and zero-padded IFFT show a similar performance. The reason is that TV works well for compressed sensing style sampling (when we have sparse samples equally from low and high frequencies) but seems to perform poorly for super-resolution (when mostly higher frequencies are missed).

To overcome TV limitations in super-resolution reconstruction, the developed edge-guided weighted-TV method (WTV) uses the complementary high frequency edge information from the structural MR modalities that are provided in higher spatial resolution than the diffusion-weighted images. A weight map is generated from the edge information and is incorporated into the optimization problem defined in equation (4.31). As shown by presented quantitative results, the developed WTV method demonstrates a significant improvement over the other two approaches (standard TV and zero-padded IFFT) in most regions of interest. In order to compare the quantitative results, paired t-test was performed and $p-value < 0.05$ was considered significant.

Since the edge information is included in non-pure regions, we did not expect to see a significant enhancement within the pure areas. However, the optimization method [99], used to solve the minimization problem in equation (4.31), is performed on the whole image; therefore, the incorporated weight map caused improvement within the whole spatial domain including pure and non-pure regions.

To investigate if the enhanced performance in each WM region of interest is still mainly due to the enhancement in their non-pure regions, we performed a correlation analysis by running linear regressions between the whole-region results

versus the results within only pure/non-pure regions. We calculated the coefficient of determination (R^2) to measure the goodness of fit to evaluate the regressions.

We computed the enhancement gained by the developed WTV over the standard TV as:

$$\Delta error = \left| W\bar{T}V_{error} - T\bar{V}_{error} \right| \tag{4.40}$$

Figures 4.14, 4.15, 4.16, and 4.17 show the scatter plots of the $\Delta error$ in each WM region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green).

A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. Here, the response variable is the $\Delta error$ in the whole white matter region of interest, and the predictor variable is the $\Delta error$ in either corresponding pure or non-pure regions.

Linear regression models the relation between a dependent, or response, variable y and one or more independent, or predictor, variables $x_1, ..., x_n$. Simple linear regression considers only one independent variable using the relation:

$$y = \beta_0 + \beta_1 x + \epsilon \tag{4.41}$$

Where β_0 is the y-intercept, β_1 is the slope (or regression coefficient), and ϵ is the error term.

Starting with a set of n observed values of x and y given by (x_1, y_1), (x_2, y_2), ..., (x_n, y_n), we form a system of linear equations using the simple linear regression relation. Following represents these equations in matrix form:

$$\begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ . & . \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \tag{4.42}$$

Let:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix}, X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ . & . \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix}, B = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \tag{4.43}$$

Now the parameters of regression (β_0 and β_1) are found by solving $Y = XB$ relation.

Then, the coefficient of determination (R^2) was calculated to measure the goodness of fit. It is a measure that allows us to determine how certain one can be in making predictions from a certain model. R^2 value falls between 0 and 1, and a closer value to 1 shows higher linear correlation between the predictor and response variables:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4.44}$$

Where $\hat{y}_i$ represents the estimated values of y_i from the calculated least-squares fit, and $\bar{y}$ is the mean of y variable. A higher coefficient of determination (R^2) for the non-pure regions, as shown in figures 4.14, 4.15, 4.16, and 4.17, demonstrates that the improvement of mean error in each WM region of interest (frontal, occipital, partial, and temporal) is highly correlated to the improvement of mean error in their corresponding non-pure areas. This is consistent with our expectation that enhancements caused by the developed WTV approach in each region of interest is mainly due to the enhancements in non-pure samples.

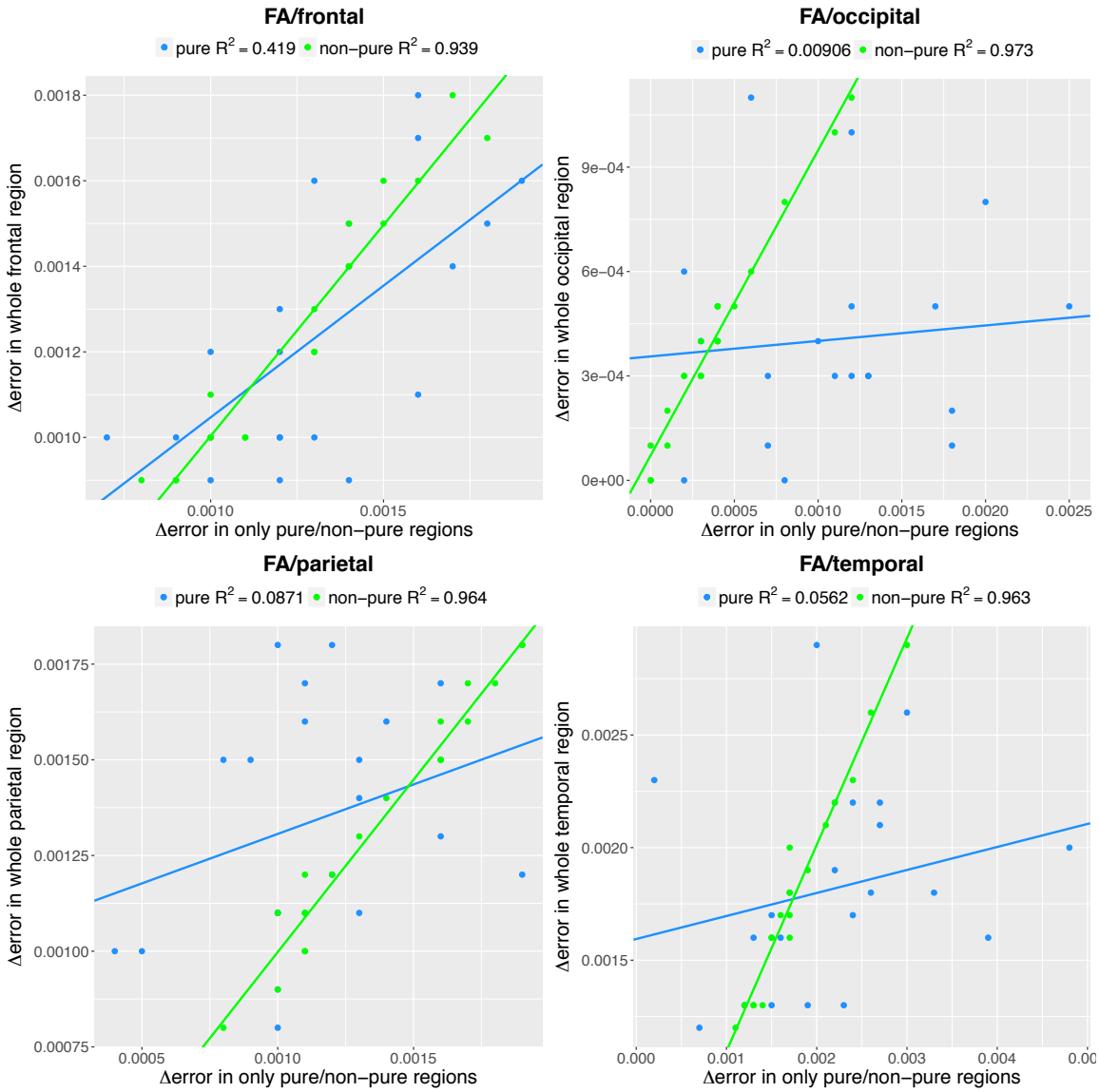


Figure 4.14: Scatter plots of $\Delta error$ of FA in each WM region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit.

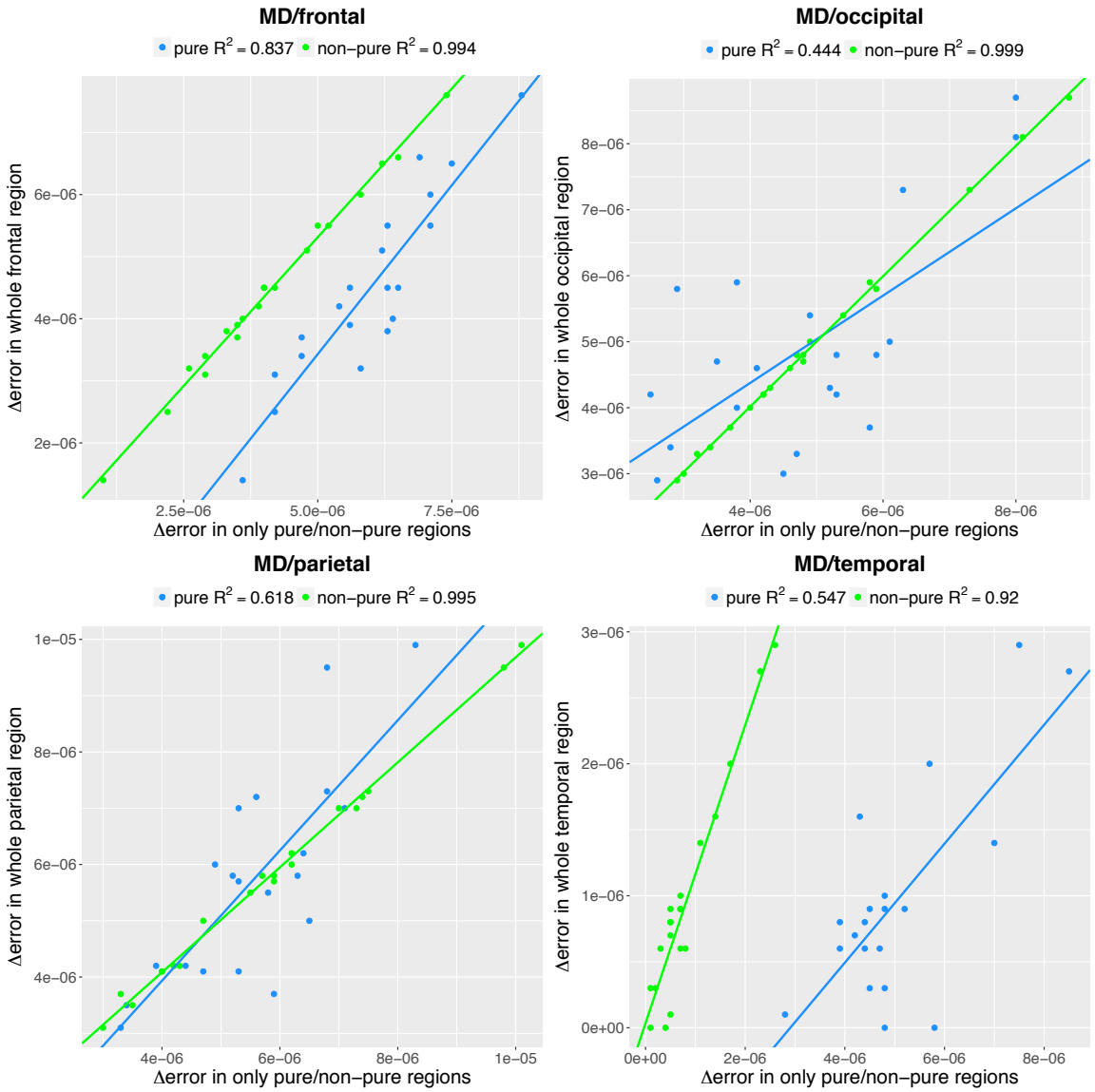


Figure 4.15: Scatter plots of $\Delta error$ of MD in each WM region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit.

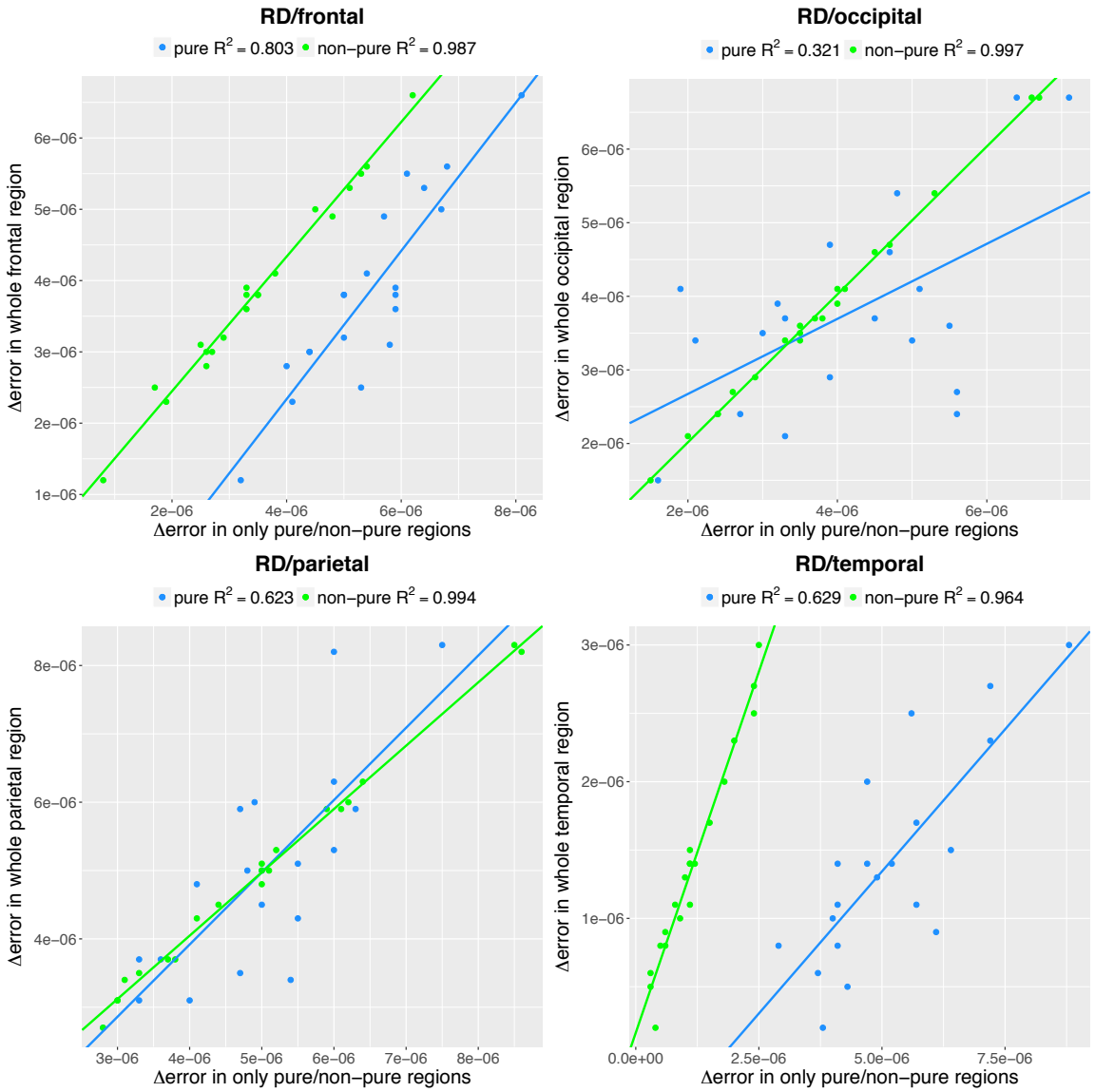


Figure 4.16: Scatter plots of $\Delta error$ of RD in each WM region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit.

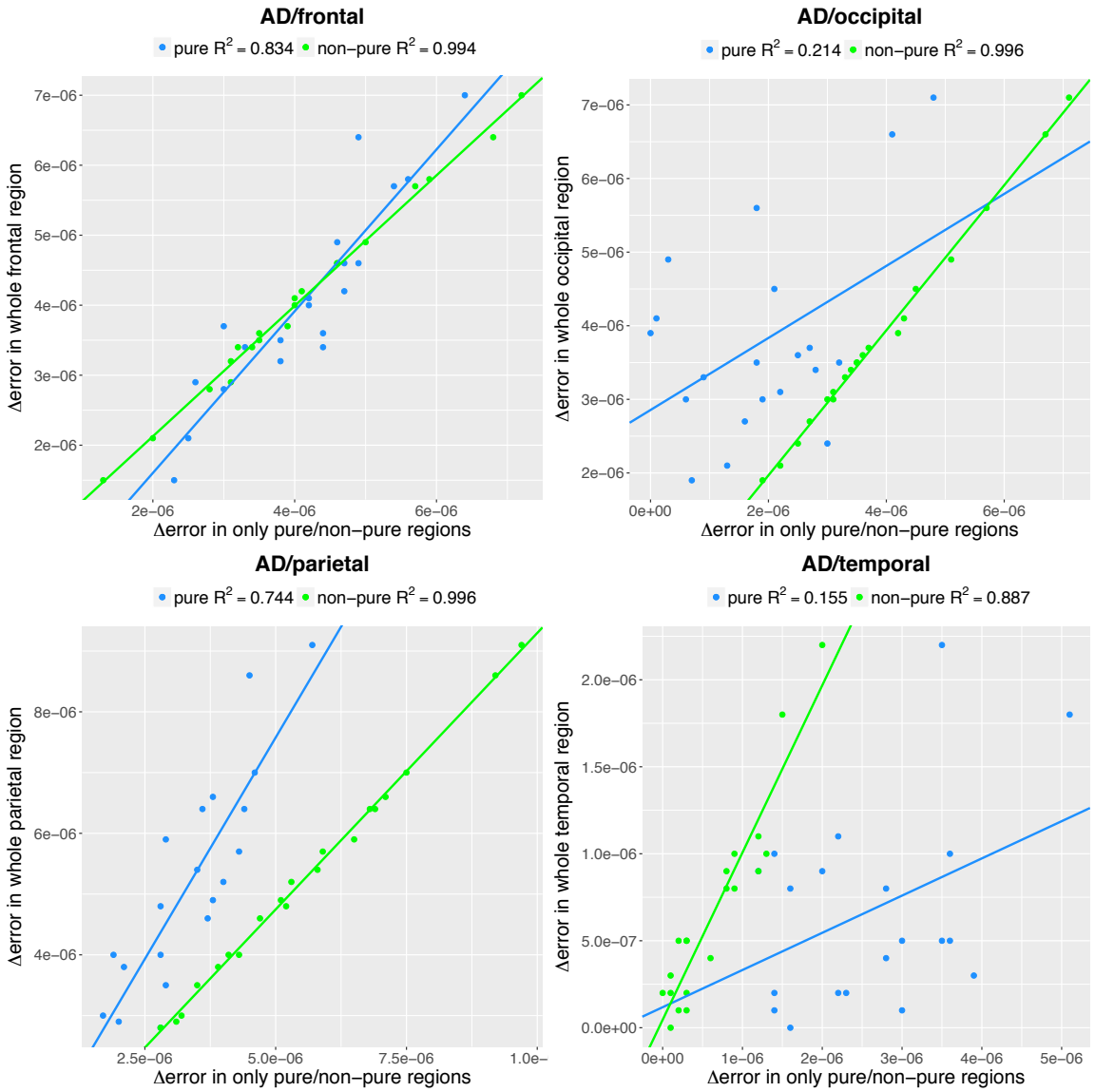


Figure 4.17: Scatter plots of $\Delta error$ of AD in each WM region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit.

4.5.4 **Approach 3**: Evaluation Based on the Differences Between Tensor Proprieties

We compared the tensor properties of the reconstructed images to the tensor properties of the HR baseline image for each test subject within four white matter regions of interest selected in four different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered. The comparison is run within the pure and non-pure regions as well, where pure and non-pure samples are computed from a pure plugs mask that is generated from the low-resolution average $b0$ image and the structural $T1$ and $T2$-weighted MR scans as described in chapter 3.

The third evaluation approach is designed to complement the approach 2 by considering diffusion orientation that is not captured by RIS measures. In this study we used Frobenius, Riemannian, and Kullback-Leibler tensor distances, described in section 1.4.5, as the error metrics to quantitatively compare the diffusion tensors from the baseline image and each or reconstructed images. Frobenius distance is not a geodesic metric, and it is not invariant to linear changes of image coordinates, so its application is limited in tensor images registration/segmentation. However, it is still a good metric in our comparison evaluation where both baseline and reconstructed images are from the same subject and are presented in the same physical coordinates.

For each test subject, the diffusion tensor images (DTI) were estimated for each of reconstructed diffusion-weighted images (DWI) and the HR baseline DWI image. Then, for each DTI voxel, we computed the distance between the diffusion tensors from the HR baseline and each of reconstructed images. Finally, for each

Table 4.4: Error mean value of all tensor metrics within each region of interest across all subjects.

Tensor Metric	Region	WTV Mean Error	TV Mean Error	IFFT Mean Error	p-value WTV vs. TV	p-value WTV vs. IFFT
Riemannian	Frontal	1.42×10^{-1}	1.54×10^{-1}	1.54×10^{-1}	$****$	$****$
	Occipital	1.69×10^{-1}	1.80×10^{-1}	1.80×10^{-1}	$**$	$**$
	Parietal	1.54×10^{-1}	1.72×10^{-1}	1.73×10^{-1}	$****$	$****$
	Temporal	1.49×10^{-1}	1.56×10^{-1}	1.56×10^{-1}	$*$	$*$
Kullback -Leibler	Frontal	7.57×10^{-2}	8.31×10^{-2}	8.32×10^{-2}	$*$	$*$
	Occipital	8.64×10^{-2}	9.75×10^{-2}	9.76×10^{-2}	$*$	$*$
	Parietal	8.22×10^{-2}	1.22×10^{-1}	1.24×10^{-1}	$***$	$***$
	Temporal	7.76×10^{-2}	8.37×10^{-2}	8.38×10^{-2}		
Frobenius	Frontal	8.23×10^{-5}	8.81×10^{-5}	8.82×10^{-5}	$****$	$****$
	Occipital	1.02×10^{-4}	1.08×10^{-4}	1.08×10^{-4}	$*$	$*$
	Parietal	9.08×10^{-5}	9.83×10^{-5}	9.83×10^{-5}	$***$	$***$
	Temporal	8.91×10^{-5}	9.06×10^{-5}	9.06×10^{-5}		

NOTE: $*$: p-value < 0.05, $**$: p-value < 0.01, $***$: p-value < 0.001, $****$: p-value < 0.0001.

region of interest, the error metric is defined as the average distances over all voxels within that region. Figures 4.18, 4.19, and 4.20 present the box plots of average errors within each region of interest for each of super-resolution reconstruction methods. All numeric results are summarized in table 4.4, where error metrics are averaged over all test subjects.

As shown by presented quantitative results, the developed WTV method demonstrates a significant improvement over the other two methods (standard TV and zero-padded IFFT) in most regions of interest. In order to compare the quantitative results, paired t-test was performed and $p - value < 0.05$ was considered significant.

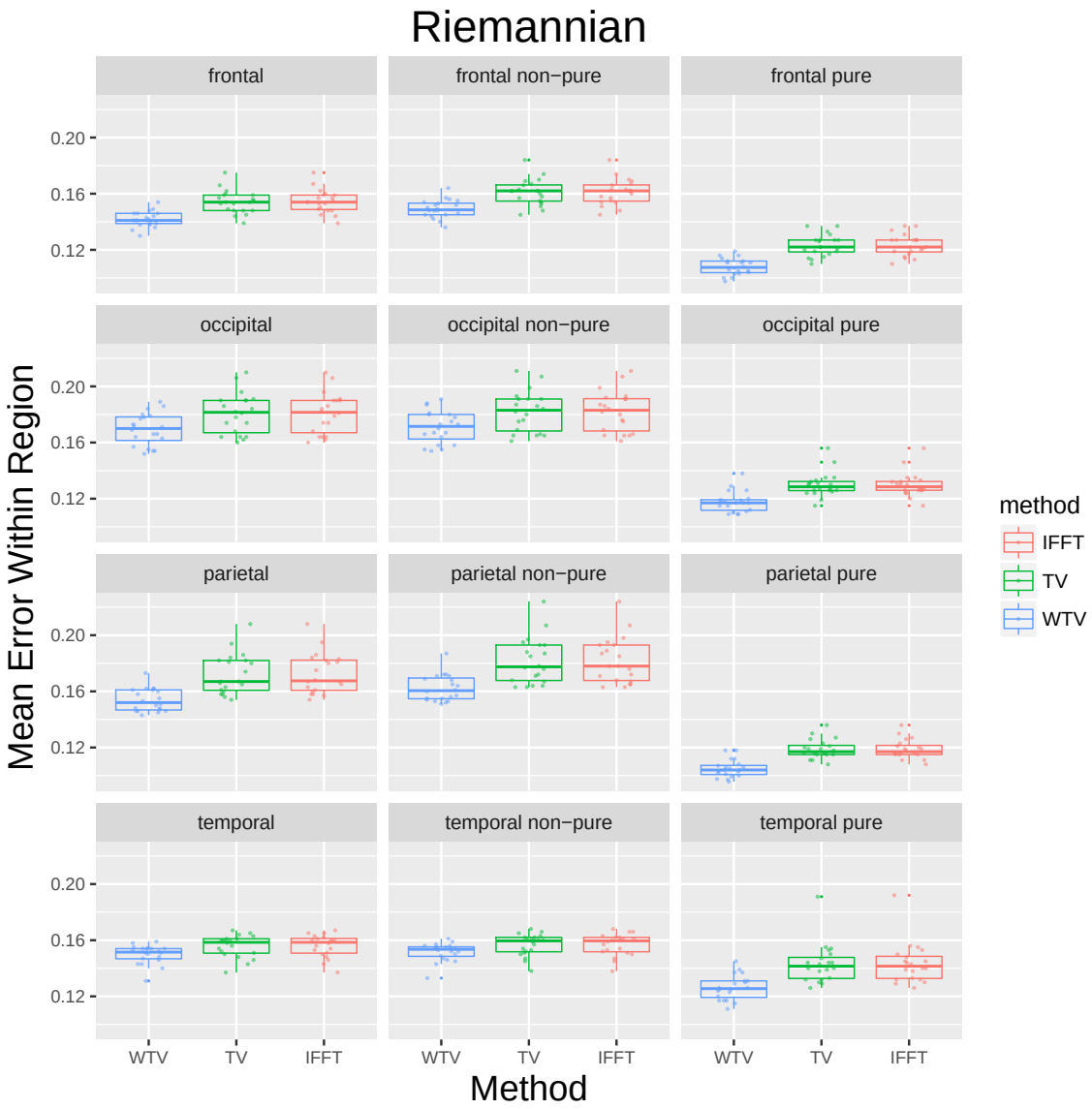


Figure 4.18: Comparing the average Riemannian error values within four WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT. The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well.

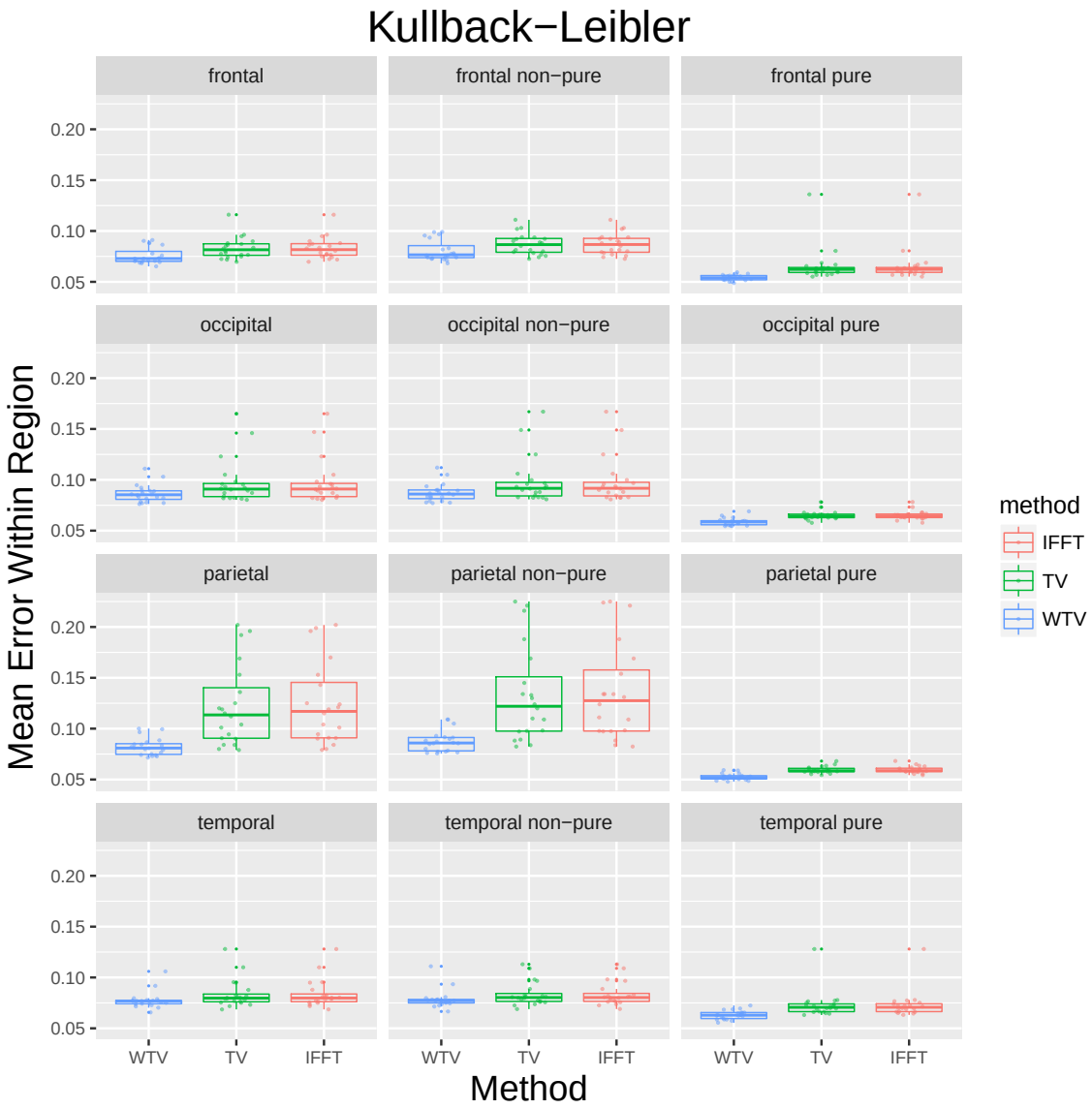


Figure 4.19: Comparing the average Kullback-Leibler error values within four WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT. The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well.

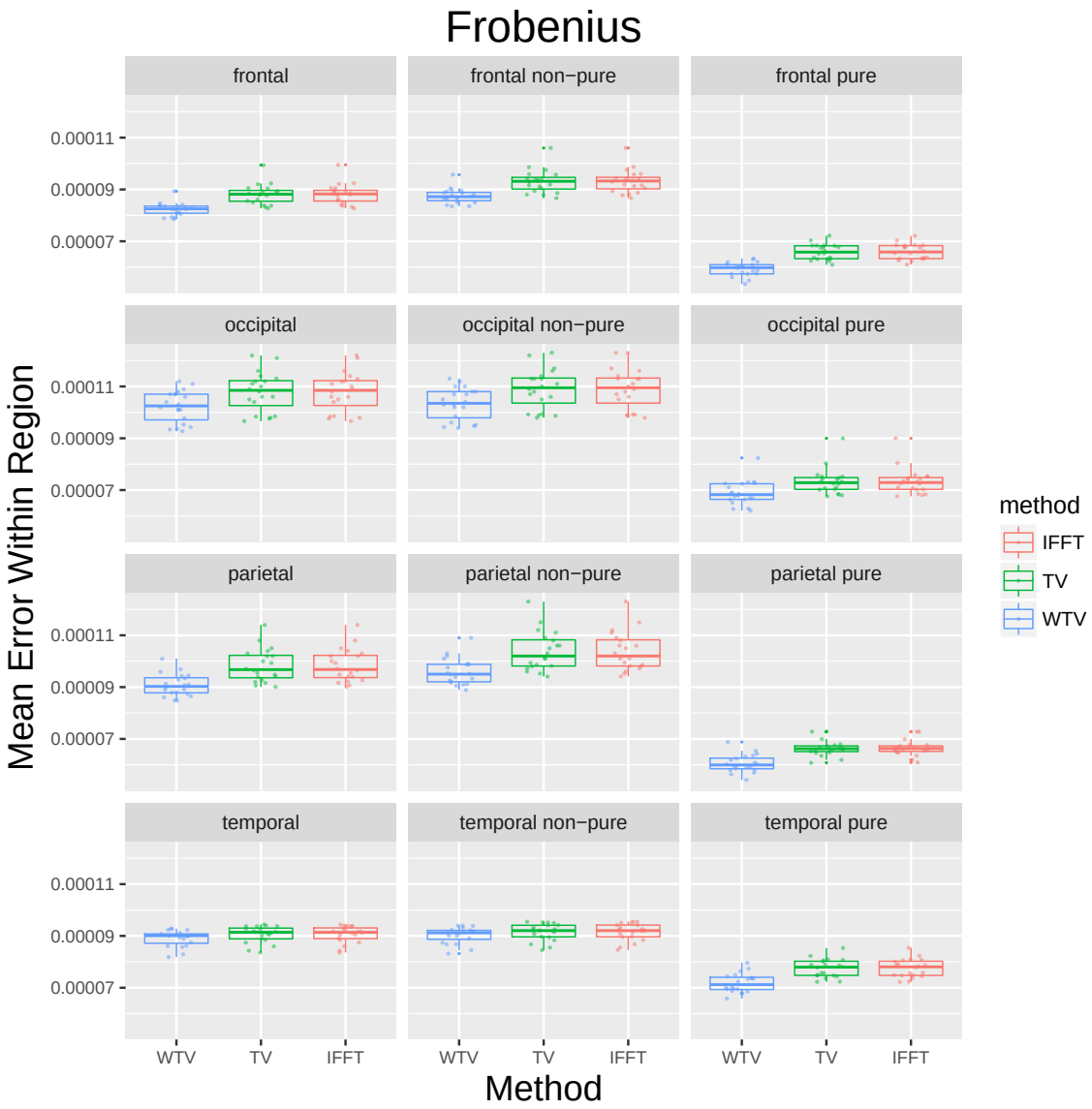


Figure 4.20: Comparing the average Frobenius norm values within four WM regions of interest between the HR baseline and each of reconstructed images by 3 different methods: Developed edge-guided weighted-TV approach (WTV), standard TV, and zero-padded IFFT. The WM regions of interests are selected in 4 different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered, and the mean of error is computed in pure and non-pure regions as well.

Same as approach 2, we performed a correlation analysis, between the whole-region results versus the results within only pure/non-pure regions, to investigate if the enhanced performance in each WM region of interest is mainly due to the enhancement in their non-pure regions. The coefficient of determination (R^2) was calculated for each linear regression to evaluate the goodness of fit. Figures 4.21, 4.22, and 4.23 show the scatter plots of the $\Delta error$, as computed in equation (4.40), in each WM region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). Also, a least-squares fit was used to fit a line between pairs of quantities and the coefficient of determination (R^2) was calculated as in equation (4.44). All figures show a higher coefficient of determination (R^2) for the non-pure regions that indicates the enhancements caused by the developed WTV approach in each region of interest is mainly due to the enhancements in non-pure samples as we expected from the method section.

4.6 Discussion and Conclusions

We introduced a novel algorithm for super-resolution reconstruction of low-resolution diffusion-weighted images using the prior anatomical information from the higher resolution structural MR modality sources. We then used three different quantitative approaches to evaluate the performance of our algorithm with two common methods in the literature: zero-padded IFFT, and standard total variation (TV).

Although TV works well for compressed sensing style sampling [113, 114, 115] (when we have sparse samples equally from low and high frequencies), it performs

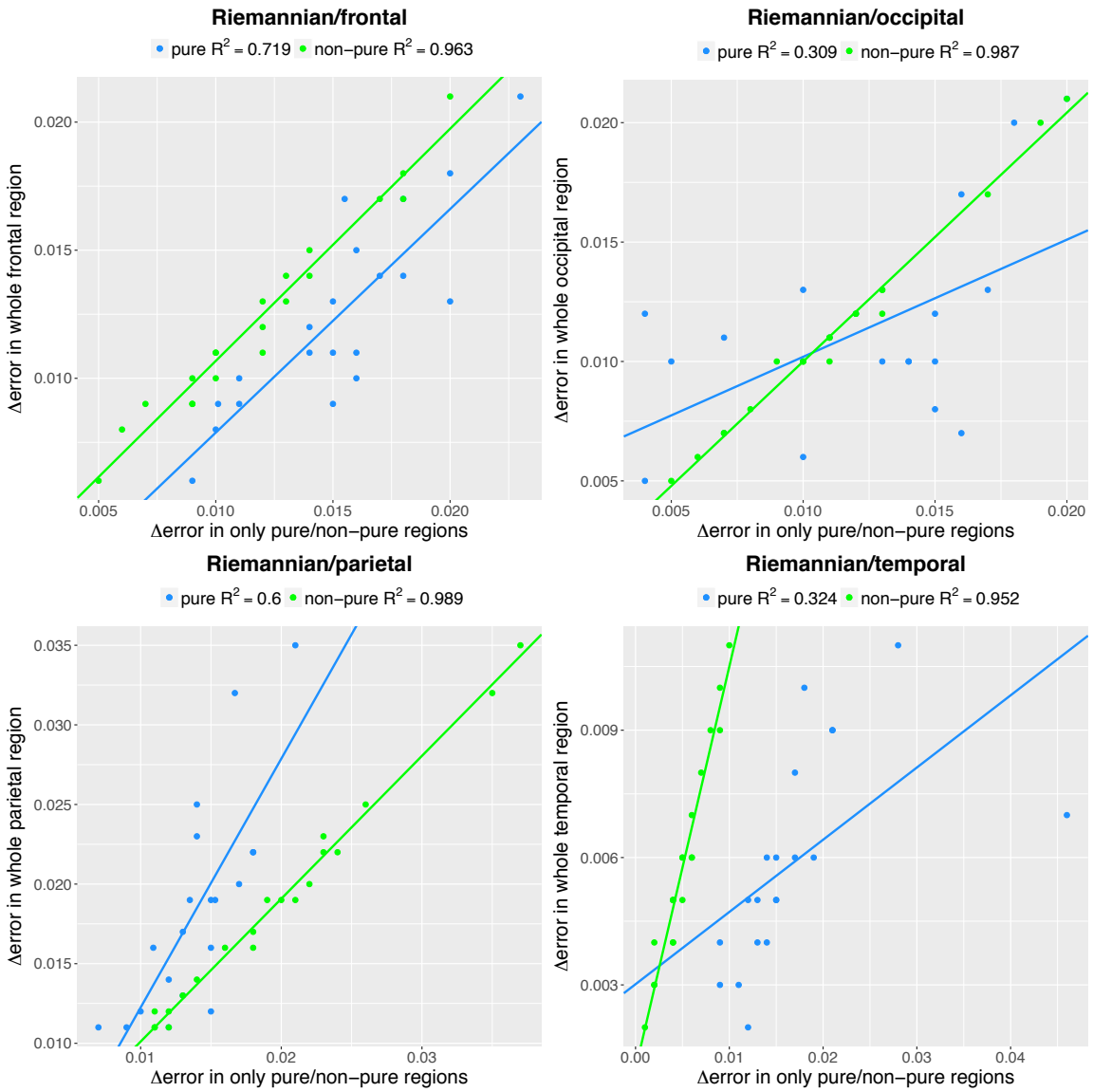


Figure 4.21: Scatter plots of $\Delta error$ of Riemannian distance in each white matter (WM) region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit.

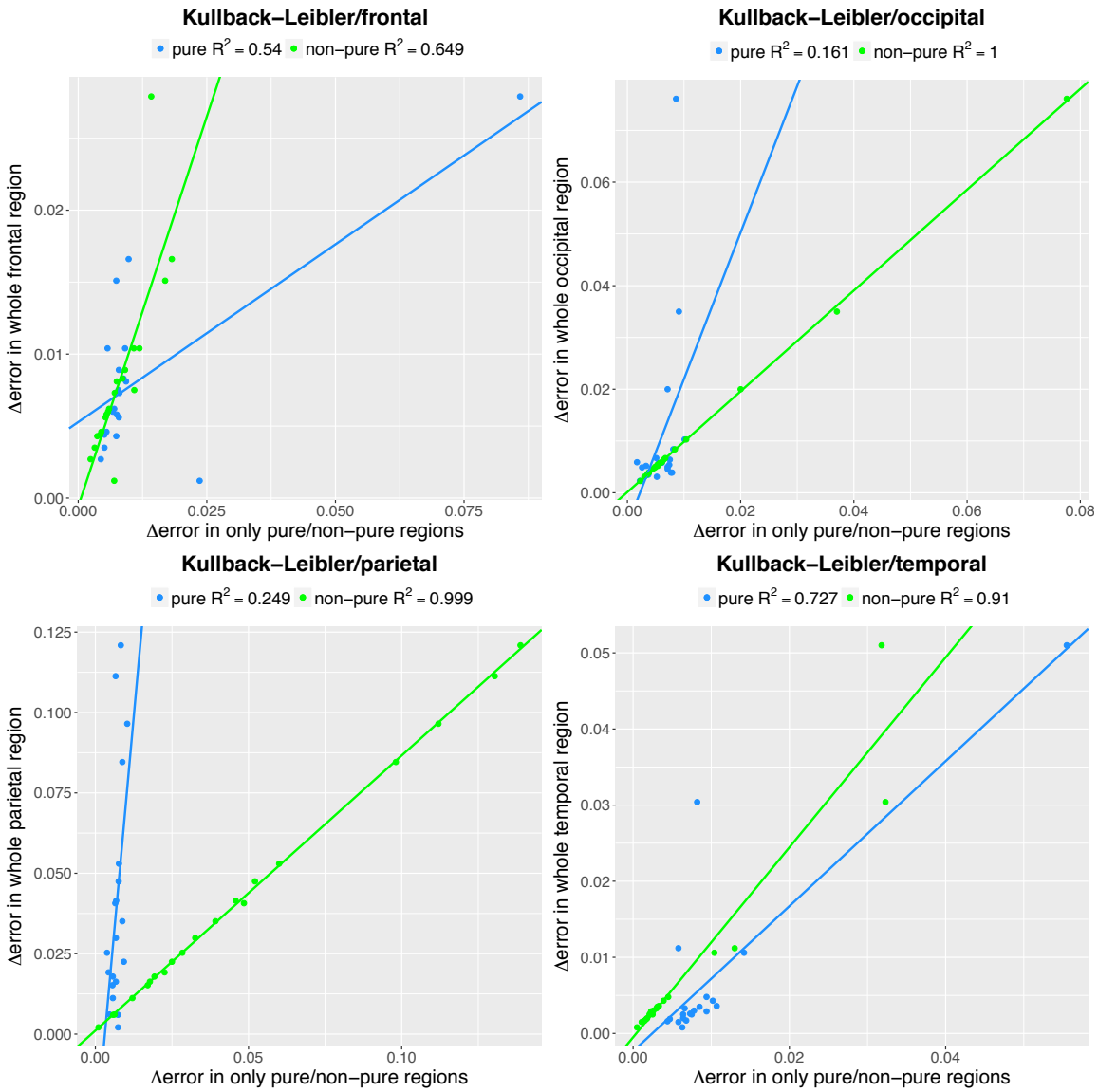


Figure 4.22: Scatter plots of $\Delta error$ of Kullback-Leibler distance in each white matter (WM) region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit.

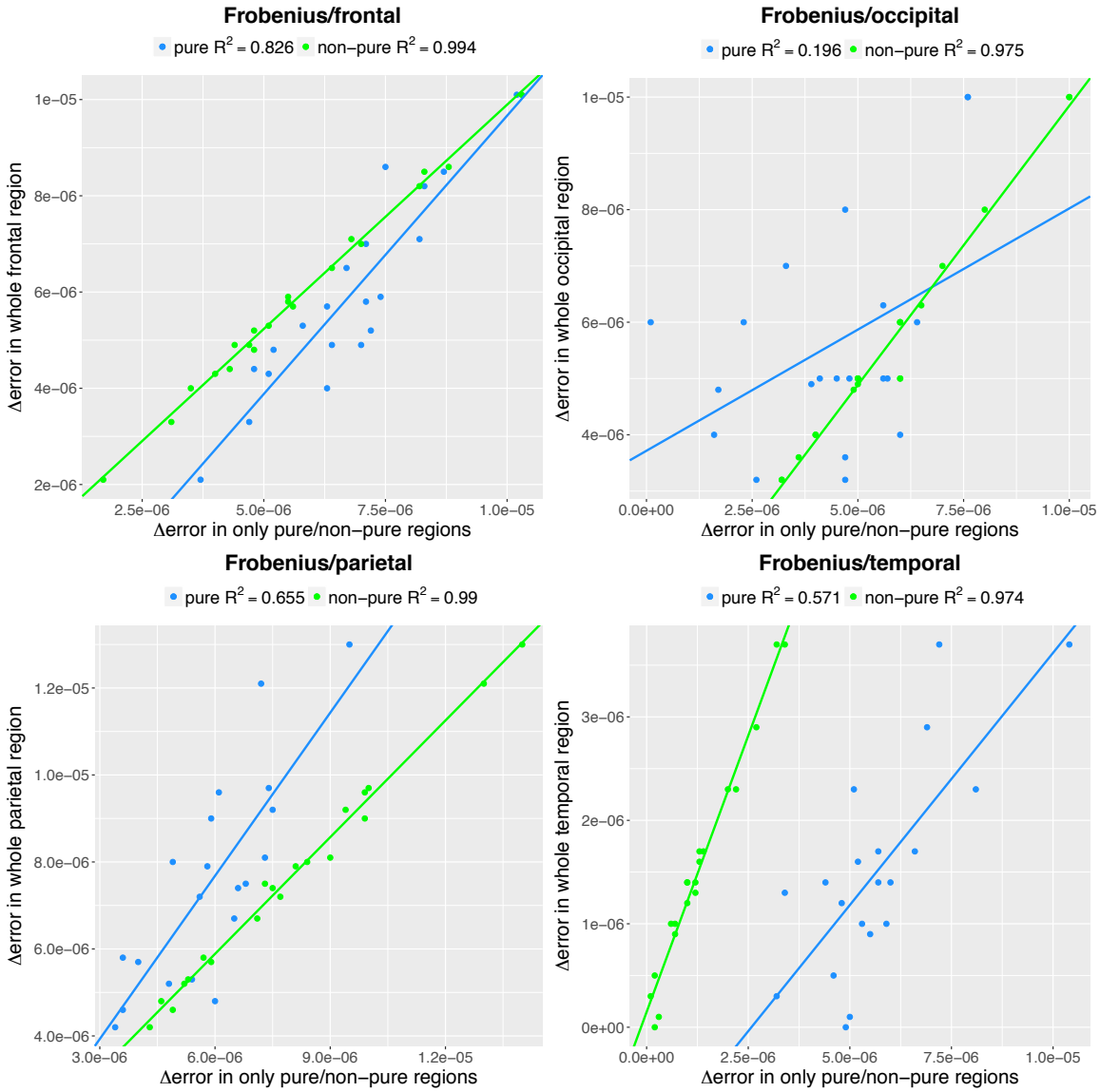


Figure 4.23: Scatter plots of $\Delta error$ of Frobenius norm in each white matter (WM) region of interest against the $\Delta error$ in only pure regions (blue) and only non-pure regions (green). A least-squares fit was used to fit a line between pairs of quantities to describe the relationship between predictor (defined in x-axis) and response (defined in y-axis) variables. The coefficient of determination (R^2) was calculated to measure the goodness of fit.

poorly for super-resolution (when mostly higher frequencies are missed) as our evaluation results showed a close performance for standard TV and zero-padded IFFT.

To overcome TV limitations in super-resolution reconstruction, we introduced an edge-guided weighted-TV method (WTV) that uses the complementary high frequency edge information from the other structural MR modalities provided in higher spatial resolution. A weight map is generated from the edge information and is incorporated into the $l1$-norm optimization process. All of our evaluation approaches, consisting of brain tractography, rotationally invariant scalars and tensor properties, demonstrated a significant improvement in the performance of developed weighted-TV algorithm over the standard TV and zero-padded IFFT.

Based on the definition of generated weight map, we expected to see a significant improvement on the performance of developed edge-guided weighted-TV over the standard TV only in high frequency boundary regions where partial volume effects encounter. However, as the optimization process is performed in the frequency domain, the generated weight map caused enhancements through all investigated white matter (WM) regions of interest in the spatial domain including both pure and non-pure regions. To investigate if the enhanced performance in each WM region of interest is still mainly due to the enhancement in their non-pure regions, we performed a correlation analysis to compute regression between the enhancement gained by the developed WTV over the standard TV in each WM region of interest against the enhancement gained in only pure regions and only non-pure regions. The results showed a higher coefficient of determination (R^2) for the non-pure regions that demonstrates

the enhancement gained in each WM region of interest (frontal, occipital, partial, and temporal) is mainly due to the enhancement gained in their corresponding non-pure regions.

CHAPTER 5
CONCLUDING REMARKS

The main purpose of this thesis was to generate tools that benefit from complementary information available from multiple structural and diffusion-weighted MR imaging modalities to advance our interpretation of brain biological architecture. We demonstrated that an improved interpretation of data in agreement with actual anatomical definitions is achieved by identifying and addressing the partial volume effects issues that occur when complementary information were combined naïvely.

First, we identified the limitations of previously developed multi-modal tools for automated tissue classification of large-scale, multi-center data. In Chapter 2, we modeled the unique anatomical states of each subject in longitudinal degenerative studies using a non-parametric fuzzy k-Nearest Neighbor (k-NN) classifier. This model augmented the output of expectation maximization (EM), a group specific classification method that uses a *priori* knowledge for all subjects in an atlas-based approach. We emphasized that using an atlas-based method is not sufficient for large-scale, multi-center longitudinal studies, due to each subject having unique anatomical states in a longitudinal degenerative study that may not be represented by prior probability distributions. Therefore, we developed a method to build up a model for each individual subject using a k-NN classifier to complement the classification results that EM produces. Chapter 2 also demonstrated segmentation enhancements when multi-modal MR modalities (T1/T2-weighted) were acquired at the same high spatial resolution with isotropic 1 mm^3 voxel sizes. However, in many datasets provided

from scanners with 1.5 Tesla scan protocol, the T2-weighted image is acquired at lower spatial resolution than T1 (usually by a factor of 2 to 3).

In Chapter 3, we demonstrated that naïvely adding low-resolution multi-modal information can adversely affect segmentation results. We investigated and explained the reason by describing partial volume effects (PVE) issue that affect more spatial samples at tissue boundaries when multi-modal scans are presented in different spatial resolutions. Then, we developed a novel approach to deal with increased PVE issue by using only those spatial samples that are not affected by partial volume composition, termed as pure samples, for initialization/training of the classification methods. The developed novel method operates in physical spatial domain and is not limited by the constrains of different voxel lattice spaces of input modalities. Pure samples were identified by computing a binary mask, called pure plugs mask. A novel integrity metric, called Mahalanobis-weighted Euclidean distance was introduced to decide if all spatial samples within a plug area belong to one anatomical tissue type. Evaluation results proved enhanced segmentation quality for gray matter (GM) and white matter (WM) when only pure samples are used in the training or initialization of the classification methods. Additionally, we demonstrated that multi-modal classification accuracy, when second modality is low-resolution, approaches the segmentation quality when all modalities are acquired in high-resolution. This technical improvement directly improved segmentation quality for the longitudinal, multi-site, international PREDICT-HD study [1], as now it was possible to use the low-resolution T2-weighted modalities that were acquired during the first 7 years of study to en-

hance the segmentation results previously generated using only T1-weighted scans. We speculate that the developed method may help to reduce the scanning time and cost in future clinical trails by dealing with PVE issue algorithmically, as this allows the collection of a single modality in a high resolution and use the other modalities acquired in lower spatial resolutions.

Chapters 2 and 3 demonstrated improved delineation of anatomical structures in structural MRIs. Chapter 4 then suggested to use the information of anatomical structures presented in high-resolution structural MR images as *a priori* knowledge to enhance the super-resolution reconstruction (SRR) of diffusion-weighted MR modalities that are acquired in low spatial resolution due to the time constraints of clinical research. Image post-processing methods are an alternative to hardware improvements that can enhance interpretation of DWI information. Providing super-resolution DWI information that is guided by the high-resolution structural scans, a more accurate assessment of characterization of brain white matter architecture and microstructure is possible. We showed the limitations of standard total variation (TV) based reconstruction approach, and introduced a novel edge-guided weighted-TV method that incorporates complementary high-resolution information from structural MR modalities into the $l1$-norm optimization process. Our evaluation approaches demonstrated the superior performance of the developed super-resolution reconstruction method in providing an accurate assessment of brain connectivity. The developed method showed better performance in recovering the high-resolution rotationally invariant scalar (RIS) measurements and high-resolution diffusion tensor properties in four

white matter regions of interest selected in four different lobes of brain (frontal, parietal, temporal and occipital) where partial volume encountered.

In conclusion, this doctoral work provided algorithmic developments for appropriate integration of complementary multi-modal information in MR studies to increase the sensitivity of volumetric and diffusion measures used by clinicians. This can lead to improved clinical and observational trials for testing therapies that may slow the progression of disease. Increasing the sensitivity of measures can lead to a substantial decrease in the number of samples needed for longitudinal and cross-sectional analysis that may reduce the cost of future studies.

APPENDIX: ABBREVIATIONS

MRI - Magnetic Resonance Imaging

DWI - Diffusion Weighted Imaging

DTI - Diffusion Tensor Imaging

DT - Diffusion Tensor

RIS - Rotationally Invariant Scalars

FA - Fractional Anisotropy

MD - Mean Diffusivity

RD - Radial Diffusivity

AD - Axial Diffusivity

ADC - Apparent Diffusion Coefficient

GM - Gray Matter

WM - White Matter

CSF - Cerebrospinal Fluid

VB - Venous Blood

PVE - Partial Volume Effect

PVC - Partial Volume Coefficient

RV - Random Variable

EM - Expectation Maximization

TPM - Tissue Probability Maps

KNN - K-Nearest Neighbors algorithm

MD - Mahalanobis distance

SRR - Super Resolution Reconstruction

HR - High Resolution

LR - Low Resolution

TV - Total Variation

WTV - Weighted Total Variation

SNR - Signal to Noise Ratio

HCP - Human Connectome Project

BC - Bhattacharyya Coefficient

AF - Arcuate Fascicle

CST - Cortico-spinal tract

REFERENCES

[1] "PREDICT-HD; an observational study of the earliest signs of huntington disease," https://www.predict-hd.net.

[2] J. P. Hornak, "The basics of MRI, 2008," *URL http://www. cis. rit. edu/htbooks/mri/index. html*, vol. 68, 2004.

[3] R. A. Pooley, "Fundamental physics of MR imaging 1," *Radiographics*, vol. 25, no. 4, pp. 1087–1099, 2005.

[4] M. A. Jacobs, T. S. Ibrahim, and R. Ouwerkerk, "MR imaging: Brief overview and emerging applications 1," *Radiographics*, vol. 27, no. 4, pp. 1213–1229, 2007.

[5] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional magnetic resonance imaging*. Sinauer Associates Sunderland, 2004, vol. 1.

[6] P. S. Tofts, G. Brix, D. L. Buckley, J. L. Evelhoch, E. Henderson, M. V. Knopp, H. B. Larsson, T.-Y. Lee, N. A. Mayr, G. J. Parker, *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*, vol. 10, no. 3, pp. 223–232, 1999.

[7] P. Hagmann, L. Jonasson, P. Maeder, J.-P. Thiran, V. J. Wedeen, and R. Meuli, "Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond 1," *Radiographics*, vol. 26, no. suppl_1, pp. S205–S223, 2006.

[8] E. Kim and H. Johnson, "Robust multi-site MR data processing: iterative optimization of bias correction, tissue classification, and registration," *Frontiers in neuroinformatics*, vol. 7, no. November, pp. 1–11, 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3831347/

[9] J. Zhuo and R. P. Gullapalli, "MR artifacts, safety, and quality control 1," *Radiographics*, vol. 26, no. 1, pp. 275–297, 2006.

[10] E. O. Stejskal and J. E. Tanner, "Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient," *The journal of chemical physics*, vol. 42, no. 1, pp. 288–292, 1965.

[11] E. Stejskal, "Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow," *The Journal of Chemical Physics*, vol. 43, no. 10, pp. 3597–3603, 1965.

[12] J. Graessner, "Frequently asked questions: Diffusion-weighted imaging (dwi)," *MAGNETON Flash*, pp. 6–9, 2011.

[13] P. B. Kingsley, "Introduction to diffusion tensor imaging mathematics: Part iii. tensor calculation, noise, simulations, and optimization," *Concepts in magnetic resonance part A*, vol. 28, no. 2, pp. 155–179, 2006.

[14] J. Matsui, "Development of image processing tools and procedures for analyzing multi-site longitudinal diffusion-weighted imaging studies," Ph.D. dissertation, The University of Iowa, 2014. [Online]. Available: http://ir.uiowa.edu/etd/4690/

[15] S. Skare, T. Li, B. Nordell, and M. Ingvar, "Noise considerations in the determination of diffusion tensor anisotropy," *Magnetic resonance imaging*, vol. 18, pp. 659–669, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0730725X00001533

[16] C. G. Koay, J. D. Carew, A. L. Alexander, P. J. Basser, and M. E. Meyerand, "Investigation of anomalous estimates of tensor-derived quantities in diffusion tensor imaging." *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, vol. 55, no. 4, pp. 930–6, Apr. 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16526013

[17] M. Niethammer and R. Estepar, "On diffusion tensor estimation," *... in Medicine and ...*, no. 1, pp. 2622–2625, 2006. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4462334

[18] N. J. Higham, "Computing the nearest correlation matrix–a problem from finance," *IMA Journal of Numerical Analysis*, vol. 22, no. 3, pp. 329–343, July 2002. [Online]. Available: http://imanum.oupjournals.org/cgi/doi/10.1093/imanum/22.3.329

[19] J. J. Shaffer, A. Ghayoor, J. D. Long, R. E.-Y. Kim, S. Lourens, L. J. O'Donnell, C.-F. Westin, Y. Rathi, V. Magnotta, J. S. Paulsen, *et al.*, "Longitudinal diffusion changes in prodromal and early hd: Evidence of white-matter tract deterioration," *Human Brain Mapping*, 2016.

[20] P. J. Basser and C. Pierpaoli, "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI," *Journal of Magnetic Resonance, Series B*, vol. 111, no. 3, pp. 209–219, Dec. 1996. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/8661285

[21] A. L. Alexander, J. E. Lee, M. Lazar, and A. S. Field, "Diffusion tensor imaging of the brain," *Neurotherapeutics*, vol. 4, no. 3, pp. 316–329, 2007.

[22] P. J. Basser and C. Pierpaoli, "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI," *Journal of magnetic resonance*, vol. 213, no. 2, pp. 560–570, 2011.

[23] M. Takahashi, D. B. Hackney, G. Zhang, S. L. Wehrli, A. C. Wright, W. T. O'Brien, H. Uematsu, F. W. Wehrli, and M. E. Selzer, "Magnetic resonance microimaging of intraaxonal water diffusion in live excised lamprey spinal cord," *Proceedings of the National Academy of Sciences*, vol. 99, no. 25, pp. 16 192–16 196, 2002.

[24] J. Acosta-Cabronero, G. B. Williams, G. Pengas, and P. J. Nestor, "Absolute diffusivities define the landscape of white matter degeneration in alzheimer's disease," *Brain*, vol. 133, no. 2, pp. 529–539, 2010.

[25] S.-K. Song, S.-W. Sun, W.-K. Ju, S.-J. Lin, A. H. Cross, and A. H. Neufeld, "Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia," *Neuroimage*, vol. 20, no. 3, pp. 1714–1722, 2003.

[26] S.-K. Song, S.-W. Sun, M. J. Ramsbottom, C. Chang, J. Russell, and A. H. Cross, "Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water," *Neuroimage*, vol. 17, no. 3, pp. 1429–1436, 2002.

[27] S.-K. Song, J. Yoshino, T. Q. Le, S.-J. Lin, S.-W. Sun, A. H. Cross, and R. C. Armstrong, "Demyelination increases radial diffusivity in corpus callosum of mouse brain," *Neuroimage*, vol. 26, no. 1, pp. 132–140, 2005.

[28] A. Einstein, *Investigations on the Theory of the Brownian Movement*. Courier Corporation, 1956.

[29] P. B. Kingsley and W. G. Monahan, "Selection of the optimum b factor for diffusion-weighted magnetic resonance imaging assessment of ischemic stroke," *Magnetic resonance in medicine*, vol. 51, no. 5, pp. 996–1001, 2004.

[30] T. Peeters, P. Rodrigues, A. Vilanova, and B. ter Haar Romeny, "Analysis of distance/similarity measures for diffusion tensor imaging," in *Visualization and Processing of Tensor Fields*. Springer, 2009, pp. 113–136.

[31] C.-F. Westin, S. Peled, H. Gudbjartsson, R. Kikinis, F. A. Jolesz, *et al.*, "Geometrical diffusion measures for MRI from tensor basis analysis," in *Proceedings of ISMRM*, vol. 97, 1997, p. 1742.

[32] A. Vilanova, S. Zhang, G. Kindlmann, and D. Laidlaw, "An introduction to visualization of diffusion tensor imaging and its applications," in *Visualization and Processing of Tensor Fields*. Springer, 2006, pp. 121–153.

[33] U. Ziyan, D. Tuch, and C.-F. Westin, "Segmentation of thalamic nuclei from dti using spectral clustering," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2006, pp. 807–814.

[34] U. Ziyan, D. Tuch, and C.-F. Westin, "Segmentation of thalamic nuclei from dti using spectral clustering," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*. Springer, 2006, pp. 807–814.

[35] P. Batchelor, M. Moakher, D. Atkinson, F. Calamante, and A. Connelly, "A rigorous framework for diffusion tensor calculus," *Magnetic Resonance in Medicine*, vol. 53, no. 1, pp. 221–225, 2005.

[36] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-euclidean metrics for fast and simple calculus on diffusion tensors," *Magnetic resonance in medicine*, vol. 56, no. 2, pp. 411–421, 2006.

[37] X. Pennec, P. Fillard, and N. Ayache, "A riemannian framework for tensor computing," *International Journal of Computer Vision*, vol. 66, no. 1, pp. 41–66, 2006.

[38] R. Vemulapalli and D. W. Jacobs, "Riemannian metric learning for symmetric positive definite matrices," *arXiv preprint arXiv:1501.02393*, 2015.

[39] Z. Wang and B. C. Vemuri, "DTI segmentation using an information theoretic tensor dissimilarity measure," *Medical Imaging, IEEE Transactions on*, vol. 24, no. 10, pp. 1267–1277, 2005.

[40] J. Tohka, "Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: A review," *World journal of radiology*, vol. 6, no. 11, p. 855, 2014.

[41] M. A. G. Ballester, A. Zisserman, and M. Brady, "Segmentation and measurement of brain structures in MRI including confidence bounds," *Medical Image Analysis*, vol. 4, no. 3, pp. 189–200, 2000.

[42] M. A. G. Ballester, A. P. Zisserman, and M. Brady, "Estimation of the partial volume effect in MRI," *Medical Image Analysis*, vol. 6, no. 4, pp. 389–405, 2002.

[43] H. S. Choi, D. R. Haynor, and Y. Kim, "Partial volume tissue classification of multichannel magnetic resonance images-a mixel model," *IEEE Transactions on Medical Imaging*, vol. 10, no. 3, pp. 395–407, 1991.

[44] P. Santago and H. D. Gage, "Quantification of mr brain images by mixture density and partial volume modeling," *IEEE Transactions on Medical Imaging*, vol. 12, no. 3, pp. 566–574, 1993.

[45] P. Santago and H. D. Gage, "Statistical models of partial volume effect," *IEEE Transactions on Image Processing*, vol. 4, no. 11, pp. 1531–1540, 1995.

[46] M. Shidahara, C. Tsoumpas, A. Hammers, N. Boussion, D. Visvikis, T. Suhara, I. Kanno, and F. E. Turkheimer, "Functional and structural synergy for resolution recovery and partial volume correction in brain pet," *Neuroimage*, vol. 44, no. 2, pp. 340–348, 2009.

[47] E. Grecchi, J. ODoherty, M. Veronese, C. Tsoumpas, G. J. Cook, and F. E. Turkheimer, "Multimodal partial-volume correction: Application to 18f-fluoride pet/ct bone metastases studies," *Journal of Nuclear Medicine*, vol. 56, no. 9, pp. 1408–1414, 2015.

[48] N. Boussion, M. Hatt, F. Lamare, Y. Bizais, A. Turzo, C. Cheze-Le Rest, and D. Visvikis, "A multiresolution image based approach for correction of partial volume effects in emission tomography," *Physics in medicine and biology*, vol. 51, no. 7, p. 1857, 2006.

[49] H. Vrooman, F. van der Lijn, and W. Niessen, "Auto-kNN: brain tissue segmentation using automatically trained knearest-neighbor classification," in *Proceedings of the MICCAI WorkshopsThe MICCAI Grand Challenge on MR Brain Image Segmentation (MRBrainS13)*, 2013.

[50] P. Anbeek, K. L. Vincken, G. S. van Bochove, M. J. P. van Osch, and J. van der Grond, "Probabilistic segmentation of brain tissue in MR imaging," *NeuroImage*, vol. 27, no. 4, pp. 795–804, 2005.

[51] M. S. Miri, A. Ghayoor, H. J. Johnson, and M. Sonka, "Comparative study of multimodal intra-subject image registration methods on a publicly available database," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2016, pp. 97 881Z–97 881Z.

[52] A. Ghayoor, J. G. Vaidya, and H. J. Johnson, "Robust automated constellation-based landmark detection in human brain imaging," *NeuroImage*, 2017.

[53] A. Ghayoor, J. G. Vaidya, and H. J. Johnson, "Development of a novel constellation based landmark detection algorithm," in *Proc. SPIE, Med. Img. 2013: Image Processing*, vol. 8669, Orlando, Florida, 2013, p. 86693F (6pages).

[54] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain," *Medical image analysis*, vol. 12, no. 1, pp. 26–41, 2008.

[55] B. B. Avants, N. Tustison, and G. Song, "Advanced normalization tools (ANTS)," *Insight Journal*, July 2009, availabel online at: http://hdl.handle.net/10380/3113.

[56] W. M. Wells, W. L. Grimson, R. Kikinis, and F. a. Jolesz, "Adaptive segmentation of MRI data." *IEEE transactions on medical imaging*, vol. 15, no. 4, pp. 429–42, Jan. 1996. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/18215925

[57] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based bias field correction of MR images of the brain." *IEEE*

transactions on medical imaging, vol. 18, no. 10, pp. 885–96, Oct. 1999. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/10628948

[58] B. B. Avants, N. Tustison, and G. Song, "Advanced Normalization Tools (ANTS)," *Image (Rochester, N.Y.)*, pp. 1–35, 2011.

[59] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian data analysis*. Chapman & Hall/CRC Boca Raton, FL, USA, 2014, vol. 2.

[60] H. A. Vrooman, C. A. Cocosco, F. van der Lijn, R. Stokking, M. A. Ikram, M. W. Vernooij, M. M. B. Breteler, and W. J. Niessen, "Multi-spectral brain tissue segmentation using automatically trained k-nearest-neighbor classification," *NeuroImage*, vol. 37, no. 1, pp. 71–81, 2007.

[61] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM*, vol. 18, no. 9, pp. 509–517, 1975.

[62] C. A. Cocosco, A. P. Zijdenbos, and A. C. Evans, "A fully automatic and robust brain MRI tissue classification method," *Medical Image Analysis*, vol. 7, no. 4, pp. 513–527, 2003.

[63] C. Cocosco, V. Kollokian, R. Kwan, and A. Evans, "Brainweb: Online interface to a 3D MRI simulated brain database," *NeuroImage*, vol. 5, p. S425, 1997.

[64] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. R. Aylward, j. V. Miller, S. Pieper, and R. Kikinis, "3D slicer as an image computing platform for the quantitative imaging network," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323–1341, Nov. 2012.

[65] H. J. Johnson, M. M. McCormick, and L. Ibanez, *The ITK Software Guide Book 1: Introduction and Development Guidelines Fourth Edition Updated for ITK version 4.7*. Kitware, Inc., 2015.

[66] H. J. Johnson, M. M. McCormick, and L. Ibanez, *The ITK Software Guide Book 2: Design and Functionality Fourth Edition Updated for ITK version 4.7*. Kitware, Inc., 2015.

[67] H. J. Johnson, A. Ghayoor, and R. E. Kim, "BRAINSTools; a suite of tools for medical image processing focused on brain analysis," https://github.com/BRAINSia/BRAINSTools, online; accessed 15-March-2017.

[68] E. Y. Kim, V. a. Magnotta, D. Liu, and H. J. Johnson, "Stable Atlas-based Mapped Prior (STAMP) machine-learning segmentation for multicenter large-scale MRI data," *Magnetic Resonance Imaging*, vol. 32, no. 7, pp. 832–844, 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/24818817

[69] R. K. Pierson, H. J. Johnson, G. Harris, H. Keefe, J. S. Paulsen, N. C. Andreasen, and V. a. Magnotta, "Fully automated analysis using BRAINS: AutoWorkup," *NeuroImage*, vol. 54, no. 1, pp. 328–336, 2011.

[70] M. Dubuisson and A. Jain, "A modified hausdorff distance for object matching," *12th IAPR International Conference*, vol. 1, pp. 566–568, 1994.

[71] A. Ghayoor, J. S. Paulsen, R. E. Kim, and H. J. Johnson, "Tissue classification of large-scale multi-site MR data using fuzzy k-nearest neighbor method," in *Proc. SPIE, Med. Img. 2016: Image Processing*, vol. 9784, San Diego, California, 2016, pp. 9784–66 (7pages).

[72] I. B. Malone, K. K. Leung, S. Clegg, J. Barnes, J. L. Whitwell, J. Ashburner, N. C. Fox, and G. R. Ridgway, "Accurate automatic estimation of total intracranial volume: a nuisance variable with less nuisance," *Neuroimage*, vol. 104, pp. 366–372, 2015.

[73] D. Liu, H. J. Johnson, J. D. Long, V. A. Magnotta, and J. S. Paulsen, "The power-proportion method for intracranial volume correction in volumetric imaging analysis," *Frontiers in neuroscience*, vol. 8, 2014.

[74] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 6, pp. 679–698, 1986.

[75] P. Filzmoser, "A multivariate outlier detection method," *in Proceedings of the Seventh International Conference on Computer Data Analysis and Modeling*, vol. 1, pp. 18–22, 2004.

[76] R. W. Brown, Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan, *Magnetic resonance imaging: physical principles and sequence design*. John Wiley & Sons, 2014.

[77] B. Scherrer, A. Gholipour, and S. K. Warfield, "Super-resolution reconstruction to increase the spatial resolution of diffusion weighted images from orthogonal anisotropic acquisitions," *Medical image analysis*, vol. 16, no. 7, pp. 1465–1476, 2012.

[78] L. Ning, K. Setsompop, O. Michailovich, N. Makris, M. E. Shenton, C.-F. Westin, and Y. Rathi, "A joint compressed-sensing and super-resolution approach for very high-resolution diffusion imaging," *NeuroImage*, vol. 125, pp. 386–400, 2016.

[79] D. C. Alexander, D. Zikic, J. Zhang, H. Zhang, and A. Criminisi, "Image quality transfer via random forest regression: applications in diffusion MRI," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*. Springer, 2014, pp. 225–232.

[80] J. Tarquino, A. Rueda, and E. Romero, "Shearlet-based sparse representation for super-resolution in diffusion weighted imaging (dwi)," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 3897–3900.

[81] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, "Super-resolution reconstruction of diffusion- weighted images using 4d low-rank and total variation," in *MICCAI 2015*, 2015.

[82] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE transactions on image processing*, vol. 10, no. 10, pp. 1521–1527, 2001.

[83] A. Yendiki, M. Reuter, P. Wilkens, H. D. Rosas, and B. Fischl, "Joint reconstruction of white-matter pathways from longitudinal diffusion mri data with anatomical priors," *NeuroImage*, vol. 127, pp. 277–286, 2016.

[84] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *Image Processing, IEEE Transactions on*, vol. 4, no. 7, pp. 932–946, 1995.

[85] A. N. Tikhonov, "On the stability of inverse problems," in *Dokl. Akad. Nauk SSSR*, vol. 39, No. 5, 1943, pp. 195–198.

[86] A. N. Tikhonov and V. I. Arsenin, *Solutions of ill-posed problems*. Vh Winston, 1977.

[87] C. Groetsch, "The theory of tikhonov regularization for fredholm equations," *104p, Boston Pitman Publication*, 1984.

[88] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[89] A. Chambolle, V. Caselles, D. Cremers, M. Novaga, and T. Pock, "An introduction to total variation for image analysis," *Theoretical foundations and numerical methods for sparse recovery*, vol. 9, no. 263-340, p. 227, 2010.

[90] D. Strong and T. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," *Inverse problems*, vol. 19, no. 6, p. S165, 2003.

[91] P. Blomgren and T. F. Chan, "Color tv: total variation methods for restoration of vector-valued images," *Image Processing, IEEE Transactions on*, vol. 7, no. 3, pp. 304–309, 1998.

[92] Y. Hu and M. Jacob, "Higher degree total variation (hdtv) regularization for image recovery," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2559–2571, 2012.

[93] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted $l1$ minimization," *Journal of Fourier analysis and applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[94] D. R. Hunter and K. Lange, "A tutorial on mm algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004.

[95] Z. Yang and M. Jacob, "Nonlocal regularization of inverse problems: a unified variational framework," *Image Processing, IEEE Transactions on*, vol. 22, no. 8, pp. 3192–3203, 2013.

[96] G. Ongie and M. Jacob, "Super-resolution MRI using finite rate of innovation curves," in *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*. IEEE, 2015, pp. 1248–1251.

[97] A. Marquina and S. J. Osher, "Image super-resolution by tv-regularization and bregman iteration," *Journal of Scientific Computing*, vol. 37, no. 3, pp. 367–382, 2008.

[98] M. Zhu and T. Chan, "An efficient primal-dual hybrid gradient algorithm for total variation image restoration," *UCLA CAM Report*, pp. 08–34, 2008.

[99] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[100] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium, *et al.*, "The wu-minn human connectome project: an overview," *Neuroimage*, vol. 80, pp. 62–79, 2013.

[101] S. N. Sotiropoulos, S. Jbabdi, J. Xu, J. L. Andersson, S. Moeller, E. J. Auerbach, M. F. Glasser, M. Hernandez, G. Sapiro, M. Jenkinson, *et al.*, "Advances in diffusion MRI acquisition and processing in the human connectome project," *Neuroimage*, vol. 80, pp. 125–143, 2013.

[102] D. Wassermann, N. Makris, Y. Rathi, M. Shenton, R. Kikinis, M. Kubicki, and C.-F. Westin, "On describing human white matter anatomy: the white matter query language," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer, 2013, pp. 647–654.

[103] M. Catani, D. K. Jones, *et al.*, "Perisylvian language networks of the human brain," *Annals of neurology*, vol. 57, no. 1, pp. 8–16, 2005.

[104] N. Makris, D. N. Kennedy, S. McInerney, A. G. Sorensen, R. Wang, V. S. Caviness, and D. N. Pandya, "Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study," *Cerebral cortex*, vol. 15, no. 6, pp. 854–869, 2005.

[105] A. Parent, *Carpenter's human neuroanatomy.* Williams & Wilkins, 1996.

[106] J. G. Malcolm, M. E. Shenton, and Y. Rathi, "Filtered multitensor tractography." *IEEE transactions on medical imaging*, vol. 29, no. 9, pp. 1664–75, Sept. 2010.

[107] C. Baumgartner, O. Pasternak, S. Bouix, C.-F. Westin, and Y. Rathi, "Filtered multi-tensor tractography using free water estimation," in *Inernational Society for Magnetic Resonance in Medicine Meeting*, 2012.

[108] A. M. Dale, B. Fischl, and M. I. Sereno, "Cortical surface-based analysis: I. segmentation and surface reconstruction," *Neuroimage*, vol. 9, no. 2, pp. 179–194, 1999.

[109] B. Fischl, "Freesurfer," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.

[110] T. Kailath, "The divergence and bhattacharyya distance measures in signal selection," *IEEE transactions on communication technology*, vol. 15, no. 1, pp. 52–60, 1967.

[111] Y. Rathi, B. Gagoski, K. Setsompop, O. Michailovich, P. E. Grant, and C.-F. Westin, "Diffusion propagator estimation from sparse measurements in a tractography framework," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 510–517, 2013.

[112] B. L. Welch, "The generalization ofstudent's' problem when several different population variances are involved," *Biometrika*, vol. 34, no. 1/2, pp. 28–35, 1947.

[113] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic resonance in medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[114] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, "An efficient algorithm for compressed MR imaging using total variation and wavelets," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on.* IEEE, 2008, pp. 1–8.

[115] C. Li, "An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing," Ph.D. dissertation, Citeseer, 2009.