TEXTURE ANALYSIS OF MULTIMODAL MAGNETIC RESONANCE IMAGES IN SUPPORT OF DIAGNOSTIC CLASSIFICATION OF CHILDHOOD BRAIN TUMOURS

SUCHADA TANTISATIRAPONG

A thesis submitted to the
University of Birmingham
for the degree of

DOCTOR OF PHILOSOPHY

School of Electronic, Electrical and Computer Engineering
College of Engineering and Physical Sciences
The University of Birmingham
August 2014

UNIVERSITY OF BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

Primary brain tumours are recognised as the most common form of solid tumours in children, with pilocytic astrocytoma, medulloblastoma and ependymoma being found most frequently. Despite their high mortality rate, early detection can be facilitated through the use of Magnetic Resonance Imaging (MRI), which is the preferred scanning technique for paediatric patients. MRI offers a variety of imaging sequences through structural and functional imaging, as well as providing complementary tissue information. However visual examination of MR images provides limited ability to characterise distinct histological types of brain tumours. In order to improve diagnostic classification, we explore the use of a computer-aided system based on texture analysis (TA) methods. TA has been applied on conventional MRI but has been less commonly studied on diffusion MRI of brain-related pathology. Furthermore, the combination of textural features derived from both imaging approaches has not yet been widely studied. In this thesis, the aim of the research is to investigate TA based on multi-centre multimodal MRI, in order to provide more comprehensive information and develop an automated processing framework for the classification of childhood brain tumours. To achieve this aim, three contributions are defined. First, a hybrid brain tumour segmentation method was developed to improve performance obtained from the use of a single segmentation technique. Tumour delineation on T2 images was found to better perform compared to FLAIR, pre-contrast T1 and post-contrast T1 based segmentation. Second, TA of single modality MRI: ADC, T2 and post-contrast T1 based classification showed improved classification outcome of childhood brain tumours compared to TA of DTI, FLAIR and pre-

contrast T1 based classification. Third, TA of multimodal MRI was found to provide potential gain in differentiating ependymomas, which are difficult to diagnose using TA of single modality of MR techniques alone.

ACKNOWLEDGEMENTS

The author would like to express sincere gratitude to the financial support from National Science and Technology Development Agency (NSTDA), Thailand and Srinakharinwirot University, Thailand. Without this scholarship, I would not have had the opportunity to study abroad.

Thank you all my supervisors: Theodoros N. Arvanitis, Andrew C. Peet and Nigel P. Davies for providing insightful suggestions and comments.

Thank you all co-authors for providing comments for publications: Nigel P. Davies, Daniel Rodriguez, Laurence Abernethy, Dorothee P. Auer, Chris. A. Clark, Richard Grundy, Tim Jaspan, Darren Hargrave, Lesley MacPherson, Martin O. Leach, Geoffrey S. Payne, Barry L. Pizer, Andrew C. Peet, Theodoros N. Arvanitis.

Thank you all at the Institute of Child Health, Birmingham Children's Hospital and colleagues for providing data and useful information: Simrandip Gill, Ben Babourina-Brooks, Ahmed Fetit, and Jane Crouch.

Thank you my parents, siblings, and my friends, particularly Asitha Mallawaarachchi and Jutharat Voraprateep for personal consultation and encouragement.

Thank you all superiors, colleagues and friends from Srinakharinwirot University for supporting my highest education.

CONTENTS

Chapter 1 Introduction 1

1.1 Motivation 1

1.2 Aim and Objectives 4

1.3 Thesis Contributions 5

1.4 Thesis Organisation 7

1.5 Summary 9

Chapter 2 Image Acquistion and Characteristics of Paediatric Brain Tumours 10

2.1 Magnetic Resonance Imaging 10

2.1.1 Nuclei in a Magnetic Field 12

2.1.2 Radio Frequency Pulse Excitation and Pulse Sequences 14

2.1.3 Gradient Encoding and Spatial Localisation 18

2.1.4 MR Image Reconstruction 20

2.1.5 Image Contrast 22

2.2 Diffusion Magnetic Resonance Imaging 24

2.2.1 The Basics of Diffusion 24

2.2.2 Diffusion Weighted Imaging 25

2.2.3 Diffusion Tensor Imaging 30

2.3 Magnetic Resonance Imaging Characteristics of Paediatric Brian Tumours 36

2.3.1 Common Paediatric Brain Tumours 36

2.3.2 Medulloblastoma 39

2.3.3 Pilocytic Astrocytoma 43

2.3.4 Ependymoma 46

2.3.5 Discussion 49

2.4 Summary 55

Chapter 3 Theoretical Background of Processing Framework 57

3.1 Introduction 57

3.2 Pre-Processing Analysis 58

3.2.1 Artefact Removal 58

3.2.2 Skull Stripping 61

3.2.3 Image Registration 63

3.2.4 Intensity Normalisation 66

3.3 Brain Tumour Segmentation 73

3.3.1 Gradient Vector Flow Snake 76

3.3.2 Multiscale Normalised Cuts Segmentation 79

3.3.3 Evaluation of Segmentation 83

3.4 Texture Analysis 87

3.4.1 Texture Analysis for the Characterisation of Childhood Brain Tumours 88

3.4.2 Review of Texture Analysis Methods 97

3.4.3 First-Order Statistics Based Textural Features 103

3.4.4 Absolute Gradient Matrix Based Textural Features 104

3.4.5 Grey-Level Co-Occurrence Matrix Based Textural Features 104

3.4.6 Grey-Level Run-Length Matrix Based Textural Features 107

3.4.7 Wavelet Transform Based Textural Features 109

3.5 Textural Feature Normalisation 115

3.6 Feature Selection 115

3.6.1 Principal Component Analysis 117

3.6.2 Minimum Redundancy and Maximum Relevance 118

3.6.3 Sequential Forward Selection 121

3.7 Classification 121

3.7.1 Support Vector Machines 122

3.7.2 Logistic Regression 124

3.8 Validation for Classification 127

3.9 Over-Fitting in Pattern Recognition 130

3.10 Statistical Measures 131

3.10.1 Sensitivity and Specificity 131

3.10.2 Hypothesis Test for Two Proportions 131

3.11 Comparative Studies 132

3.12 Summary 135

Chapter 4 Methodology 136

4.1 Introduction 136

4.2 Experiment 1: Pre-Processing Analysis for Brain Imaging Data 137

4.2.1 Artefact Removal 138

4.2.2 Skull Stripping 139

4.2.3 Diffusion Magnetic Resonance Image Reconstruction 139

4.2.4 Image Registration 139

4.2.5 Intensity Standardisation 140

4.3 Experiment 2: Brain Tumour Segmentation 141

4.3.1 Combined MNcut and GVF Snake Hybrid Method 141

4.3.2 Evaluation of Segmentation 142

4.4 Experiment 3: Classification of Brain Tumours Based on Texture Analysis of Single Modality MRI 144

4.4.1 Principal Component Analysis 146

4.4.2 Combined mRMR and SFS method 146

4.5 Experiment 4: Classification of Brain Tumours Based on Texture Analysis of Multimodal MRI 147

4.6 Dataset 147

4.7 Quality Control 148

4.8 Summary 149

Chapter 5 Experiment 1: Pre-Processing Analysis for Brain Imaging Data 151

5.1 Introduction 151

5.2 Material 152

5.3 Results and Discussion 153

5.3.1 Artefact Removal 153

5.3.2 Skull Stripping 154

5.3.3 Diffusion MRI Reconstruction 156

5.3.4 Image Registration 156

5.3.5 Intensity Normalisation 158

5.4 Conclusions 160

Chapter 6 Experiment 2: Brain Tumour Segmentation 162

6.1 Introduction 162

6.2 Material 163

6.3 Results and Discussion 164

6.3.1 Segmentation Outcomes 164

6.3.2 Five-Point Likert Scale of Evaluation 167

6.3.3 Supervised Evaluation 172

6.3.4 Constraints of the Proposed Method 174

6.3.5 Adaptation to Clinical Environment 175

6.3.6 Future Work 176

6.4 Conclusions 176

Chapter 7 Experiment 3: Classification of Brain Tumours Based on Texture Analysis of Single Modality MRI 178

7.1 Introduction 178

7.2 Material 179

7.3 Results and Discussion 180

7.3.1 PCA Based Classification 182

7.3.2 Combined mRMR and SFS Method Based Classification 184

7.3.3 Comparative Analysis 188

7.3.4 Evaluation of Classifier 192

7.3.5 Effects from Texture Analysis 193

7.3.6 Comparative Studies 197

7.4 Conclusions 198

Chapter 8 Experiment 4: Classification of Brain Tumours Based on Texture Analysis of Multimodal MRI 200

8.1 Introduction 200

8.2 Material 201

8.3 Methods 202

8.4 Results and Discussion 204

8.4.1 Examining Textural Features 205

8.4.2 Classification Outcomes 210

8.4.3 Comparative Statistical Test 213

8.4.4 Visual Analysis 215

8.4.5 Building Multinomial Logistic Regression Models 218

8.4.6 Recommended Processing Framework 223

8.4.7 Biological Perspective 225

8.4.8 Classification Difficulties and Future Work 226

8.5 Conclusions 228

Chapter 9 Conclusions 230

Appendix A: Textural Features Derived From GLCM and GLRLM 233

A.1 Grey-Level Co-Occurrence Matrix Based Textural Features 233

A.2 Grey-Level Run-Length Matrix Based Textural Features 236

Appendix B: Pre-Processing Shell Script 238

Appendix C: Imaging Parameters for Experimental Dataset 240

C.1 Imaging Parameters of Datasets for Experiment 1 240

C.2 Imaging Parameters of Datasets for Experiment 2 246

C.3 Imaging Parameters of Datasets for Experiment 3 248

C.4 Imaging Parameters of Datasets for Experiment 4 251

Appendix D: Additional Example Images **255**

Appendix E: Publication (BC ISMRM 2012) **257**

Appendix F: Publication (BioMed 2013) **259**

Appendix G: Publication (MEDICON 2013) **270**

Appendix H: Publication (ISMRM 2014) **276**

Bibliography **278**

LIST OF ABBREVIATIONS AND TERMS

Abbreviation	Full name
ADC	Apparent Diffusion Coefficient
BET	Brain Extraction Tool
BSE	Brain Surface Extractor
CCLG	Children's Cancer and Leukaemia Group
CNS	Central Nervous System
CR	Correlation Ratio
CSF	Cerebrospinal Fluid
CT	Computed Tomography
DOF	Degree of Freedom
DTI	Diffusion Tensor Imaging
DWI	Diffusion Weighed Imaging
EP	Ependymoma
EPI	Echo Planar Imaging
FA	Fractional Anisotropy
FID	Free Induction Decay
FLAIR	Fluid Attenuated Inversion Recovery
FLIRT	FMRIB's Linear Image Registration Tool
FMRIB	The Oxford Centre for Functional Magnetic Resonance Imaging of the Brain

FSL	FMRIB Software Library
GLCM	Grey-Level Co-occurrence Matrix
GLRLM	Grey-Level Run-Length Matrix
GVF	Gradient Vector Flow
HMRI	Hemodynamic Magnetic Resonance Imaging
HWA	Hybrid Watershed Algorithm
LDA	Linear Discriminant Analysis
LOOB	Leave-One-Out Bootstrap
LOOCV	Leave-One-Out Cross Validation
MAPS	Multi-Atlas Propagation and Segmentation
MB	Medulloblastoma
McStrip	Minneapolis Consensus Strip
MD	Mean Diffusivity
MDH	Modified Hausdorff Distance
MI	Mutual Information
MID	Mutual Information Difference
MIQ	Mutual Information Quotient
MLR	Multinomial Logistic Regression
mm	Millimetre
MNcut	Multiscale Normalised Cuts
MRI	Magnetic Resonance Imaging
mRMR	minimal-Redundancy-Maximal-Relevance

MRS	Magnetic Resonance Spectroscopy
MRSI	Magnetic Resonance Spectroscopy Imaging
ms	Millisecond
NMI	Normalised Mutual Information
NMR	Nuclear Magnetic Resonance
PA	Pilocytic Astrocytoma
PC	Principal Component
PCA	Principal Component Analysis
PNN	Probabilistic Neural Network
PWI	Perfusion Weighted Imaging
RF	Radio Frequency
ROI	Region of Interest
SFS	Sequential Forward Selection
SVM	Support Vector Machine
TA	Texture Analysis
TE	Echo Time
TI	Inversion Time
TR	Repetition Time
WHO	World Health Organization

LIST OF FIGURES

Figure 2.1: A simplified diagram of a MRI system. 11

Figure 2.2: Precession of a nucleus around its own static magnetic field axis. 12

Figure 2.3: Energy states and magnetisation of nuclear spins. 13

Figure 2.4: De-phasing of nuclear spins after a 90° RF pulse excitation. 14

Figure 2.5: The change in magnetisation of a spin-echo sequence 17

Figure 2.6: The change in magnetisation of an inversion recovery sequence 17

Figure 2.7: A typical gradient-echo imaging sequence. 19

Figure 2.8: A typical spin-echo imaging sequence. 20

Figure 2.9: A typical inversion recovery imaging sequence. 20

Figure 2.10: A conversion of RF signal to a digital image. 21

Figure 2.11: Complementary image contrast of brain tissue acquired based on T2, pre-contrast T1 and FLAIR image 23

Figure 2.12: Signal intensity of restricted and unrestricted space based on DWI and ADC. 27

Figure 2.13: The difference of imaging features and deterioration of SNR-based DWI at b-value = 0 and 800 s/mm^2 29

Figure 2.14: The difference of imaging features and deterioration of SNR-based DWI at b-value = 0, 500 and 1000 s/mm^2 29

Figure 2.15: Comparison of white matter on T2, pre-contrast T1 and FA images. 30

Figure 2.16: Fibre structures and shape of diffusion tensors 35

Figure 2.17: Some brain components and location on a sagittal T2-weighted image 37

Figure 2.18: Imaging features of grey matter, white matter and CSF 38

Figure 2.19: Imaging characteristics of a medulloblastoma 41

Figure 2.20: Imaging characteristics of a desmoplastic medulloblastoma 42

Figure 2.21: Comparison of two FLAIR images of medulloblastoma cases which have different signal intensity level compared to normal tissues. 42

Figure 2.22: Imaging characteristics of a typical pilocytic astrocytoma 44

Figure 2.23: Pilocytic astrocytomas with almost exclusively solid and no prominent cystic component. 45

Figure 2.24: Imaging characteristics of a typical grade II ependymoma 48

Figure 2.25: Imaging characteristics of a supratentorially located anaplastic ependymoma 49

Figure 2.26: Difficulty in differentiating the three tumour types. 54

Figure 3.1: A general framework of multimodal MRI based texture analysis for classification of childhood brain tumours. 58

Figure 3.2: Intensity normalisation based on linear normalisation method. 71

Figure 3.3: Intensity normalisation based on piecewise linear transform method 72

Figure 3.4: Three scales of the MNcut graph decomposition with R = 1. 83

Figure 3.5: Two images having the same textural features of first-order statistics. 99

Figure 3.6: An example image for calculating an absolute gradient matrix. 104

Figure 3.7: The four grey-level co-occurrence matrices of a digital image 106

Figure 3.8: The four grey-level run-length matrices of a digital image 108

Figure 3.9: A 2D wavelet decomposition with 1-scale and 2-scale decomposition. 111

Figure 3.10: Scaling and wavelet functions of Haar Wavelet 112

Figure 3.11: Scaling and wavelet functions of $D4$ 113

Figure 3.12: Scaling and wavelet functions of $sym4$..... 114

Figure 3.13: A diagram of sequential forward selection. 121

Figure 4.1: Processing pipeline of texture analysis of childhood brain tumour based on single modality and multimodal MRI. 137

Figure 4.2: Pre-processing analysis of brain imaging data based MRI. 138

Figure 4.3: The proposed segmentation algorithm 142

Figure 4.4: Classification of brain tumours based on TA of Single Modality MRI. 145

Figure 4.5: Classification of brain tumours based on TA of Multimodal MRI. 147

Figure 5.1: A comparison between original and eddy current corrected images. 154

Figure 5.2: Satisfactory skull stripping outcomes obtained from BET method. 155

Figure 5.3: Failure of skull removal on some datasets. 155

Figure 5.4: Reconstruction of DTI images 156

Figure 5.5: Failure of registration based on CR similarity measure..... 157

Figure 5.6: Successful registration for a medulloblastoma with NMI cost function 158

Figure 5.7: A comparison of intensity normalisation performed on a medulloblastoma . 159

Figure 6.1: The four MRI types of a pilocytic astrocytoma are segmented by the combined MNcut and GVF snake method. 165

Figure 6.2: Segmentation results derived from two K values of the MNcut method. 166

Figure 6.3: A segmentation evaluation form. 167

Figure 6.4: Semi-automatic segments of medulloblastoma, pilocytic astrocytoma, and ependymoma based on T2, FLAIR, pre-contrast T1, and post-contrast T1. 168

Figure 6.5: The average score with standard deviation obtained from segmentation scores rated by the four evaluators 169

Figure 6.6: A comparison of tumour boundaries derived from manual segmentation and semi-automatic segmentation based on T2 image 172

Figure 7.1: Tumour boundaries of the three tumour types. 181

Figure 7.2: The variation of classification accuracy with respective to the number of PCs for TA based on conventional and diffusion MRI. 183

Figure 7.3: Plot of 1st and 2nd PC of features selected by combined mRMR and SFS method 187

Figure 7.4: Classification accuracy of MB, PA, EP and overall tumour types obtained from PCA, combined mRMR and SFS and original features based classification 189

Figure 7.5: The variation of classification accuracy with respect to the number of grey levels for GLCM based TA. 195

Figure 7.6: The variation of classification accuracy with respect to the number of grey levels for GLRLM based TA. 195

Figure 8.1: Classification based on combined multimodal MRI based TA. 203

Figure 8.2: Classification based on additive probability outcome. 203

Figure 8.3: Classification based on superposition of weighted outcomes. 204

Figure 8.4: Data distribution of the two top-ranked features derived from ADC, T2 and post-contrast T1 dataset. 209

Figure 8.5: Classification accuracy derived from single modality and multimodal MRI based texture analysis. 212

Figure 8.6: Two cases misclassified by ADC based TA but correctly predicted by the superposition of weighted outcomes. 217

Figure 8.7: Two cases misclassified by ADC based TA but correctly predicted by the combined multimodal MRI based TA approach. 217

Figure 8.8: Four cases misclassified by the superposition of weighted outcomes. 217

Figure 8.9: Four cases misclassified by the combined multimodal MRI based TA approach. 217

Figure 8.10: The two recommended processing pipelines for the classification of childhood brain tumours. 224

LIST OF TABLES

Table 2.1: Comparison of signal intensity of DWI and ADC images. 28

Table 2.2: MRI features for medulloblastoma, pilocytic astrocytoma and ependymoma..50

Table 3.1: The common texture analysis methods applied on conventional, diffusion and multimodal MRI of brain tumours. 97

Table 3.2: Advantages and disadvantage of TA for defining brain tumours. 102

Table 3.3: List of 20 textural features derived from GLCM. 107

Table 3.4: List of 11 textural features derived from GLRLM. 109

Table 4.1: Meaning of scores for the five-point Likert scale evaluation form. 143

Table 5.1: Number of MRI and diffusion MRI dataset of three tumour types. 152

Table 5.2: Imaging parameters for experiment 1. 153

Table 5.3: Mean and standard deviation of intensity for each tumour and MRI type. 160

Table 6.1: Imaging parameters for experiment 2. 163

Table 6.2: Comparison between manual segmentation and semi-automatic segmentation. 173

Table 7.1: Number of subjects for each tumour type for experiment 3. 180

Table 7.2: Imaging parameters for experiment 3. 180

Table 7.3: Sensitivity and specificity for medulloblastoma, pilocytic astrocytoma, ependymoma at the optimal number of principal components. 183

Table 7.4: Textural features selected by using mRMR and followed by SFS method. 185

Table 7.5: Sensitivity and specificity for medulloblastoma, pilocytic astrocytoma, ependymoma derived from the combined mRMR and SFS method based classification.186

Table 7.6: P-values of statistical hypothesis test between the classification accuracy obtained from three pairwise comparisons. 190

Table 7.7: Comparison of classification error estimated from LOOCV, LOOB, .632 estimator and .632+ estimator. 191

Table 7.8: P-values of statistical hypothesis test between the classification accuracy obtained from two MR-based image types. 192

Table 7.9: The optimal number of grey levels that gives the highest classification accuracy derived from GLCM and GLRLM. 196

Table 7.10: Classification accuracy obtained from the optimal number of grey levels vs nine grey levels. 196

Table 7.11: Sensitivity and specificity derived from T2 and pre-contrast T1 based TA, reported in Vlachou et al. (19) and our study. 197

Table 7.12: Comparison of classification accuracy derived from ADC based TA, reported in Rodriguez Gutierrez et al. (30) and our study. 198

Table 8.1: Number of subjects for each tumour type for experiment 4. 202

Table 8.2: Imaging parameters for experiment 4. 202

Table 8.3: Classification accuracy and deviance of the ten top-ranked textural features selected by using mRMR method from ADC dataset. 206

Table 8.4: Classification accuracy and deviance of the ten top-ranked textural features selected by using mRMR method from T2 dataset. 206

Table 8.5: Classification accuracy and deviance of the ten top-ranked textural features selected by using mRMR method from post-contrast T1 dataset. 207

Table 8.6: The optimal feature set for ADC, T2, post-contrast T1 and the combined multimodal MRI based TA. 208

Table 8.7: Confusion matrix for ADC dataset. 211

Table 8.8: Confusion matrix for T2 dataset. 211

Table 8.9: Confusion matrix for post-contrast T1 dataset..... 211

Table 8.10: Confusion matrix of the additive probability outcome. 212

Table 8.11: Confusion matrix of the combined multimodal MRI based TA. 212

Table 8.12: Confusion matrix of the superposition of weighted outcomes..... 212

Table 8.13: P-values of statistical hypothesis test of pairwise classification comparison. 213

Table 8.14: Beta coefficients and goodness of fit derived from ADC dataset..... 219

Table 8.15: Beta coefficients and goodness of fit derived from T2 dataset..... 220

Table 8.16: Beta coefficients and goodness of fit derived from post-contrast T1 dataset. 221

Table 8.17: Beta coefficients and goodness of fit derived from the combined multimodal MRI based TA approach..... 222

Table 8.18: Classification accuracy based on the analysis of radiological reports and our study..... 223

LIST OF SYMBOLS AND NOTATION

Symbol	Notation
$\mathcal{R}$	The set of real numbers.
$\mathcal{R}^C$	The set of C real numbers.
i	The imaginary unit i.e. $\sqrt{-1}$.
x^T	The transpose of element x.
x'	The first derivation of element x.
x''	The second derivation of element x.
∇	Gradient operator.
∇^2	Laplacian operator.
$A = \{a_1, a_2, \dots, a_m\}$	The set of A has m number of element a.
$\lvert A \rvert$	The number of element of set A.
$\lVert A \rVert$	The norm of the element A.
$a \in A$	The element a is a member of the set A.
$max(A)$	The maximum value of the set of A.
$min(A)$	The minimum value of the set of A.

Lowercase or uppercase symbols denote scalar element, e.g. a, b, A and B.

Lowercase boldface or uppercase boldface with a hat symbols denote vectors, e.g. $\boldsymbol{b}$ and $\widehat{\boldsymbol{B}}$.

Uppercase boldface symbols denote matrices, e.g. $\boldsymbol{P}$.

CHAPTER 1 INTRODUCTION

1.1 Motivation

Brain tumours can be categorised as either primary or secondary. Primary brain tumours develop from normal cells in the brain, whereas secondary brain tumours arise when cancer cells from other parts of the body spread to the brain. Brain tumours can grow, infiltrate, become more malignant, as well as result in morbidity and mortality. Furthermore, they can have diversity of biology and clinical behaviour. In order to provide an accurate diagnosis and treatment, it is of prime importance to understand their pathology. The pathological behaviour of Central Nervous System (CNS) tumours has been defined by the World Health Organization (WHO) with two of the most important classification criteria: level of aggressiveness and cell of origin. Level of aggressiveness divides tumours into a four grade system (1). Grade I represents the biologically least aggressive tumours (benign or non-cancerous) and grade IV denotes the biologically most aggressive tumours (malignant or cancerous). The cell of origin can be examined by the histopathology method which performs surgical extraction of specimen tissue and examines it in microscopic detail.

Primary brain tumours are the second most common group of cancers after leukaemia and the most common form of solid tumours in children (2). They account for approximately 20-25% of all primary paediatric malignancies (3). Although cancer is found only around 1 in 500 in children, brain and CNS tumours are the most common cause of cancer deaths, accounting for about one third of all cancer deaths in children in the

United Kingdom (4). The majority of primary childhood brain tumours have been reported to be three types of posterior fossa tumours: pilocytic astrocytoma (PA) - grade I, medulloblastoma (MB) - grade IV and ependymoma (EP) - grade I-III (5). These three tumour types are not commonly found in adults and even when they occur, their prognosis in adults is not necessarily the same as in children (6-8).

Imaging is an essential non-invasive tool in clinical diagnosis and management of brain tumours. The two most popular imaging techniques are computed tomography (CT) and magnetic resonance imaging (MRI). CT imaging is broadly accessible in most clinical centres and, as a result, is first performed in paediatric patients with presentation of suspected neurological symptoms (9). Compared with MRI, CT provides better image contrast of small amounts of calcification within tumours (10), requires shorter scanning time and is more cost effective (11). Although CT offers several advantages, MRI is a preferred modality in brain tumour imaging because of several benefits. MRI can offer better sensitivity of tumour delineation because of its higher soft tissue contrast (12). It also provides improved diagnosis of paediatric brain tumours because of higher sensitive to the detection of small tumours, particularly those in the posterior fossa region (9, 13).

Furthermore, MRI provides complementary information from numerous imaging sequences, through the use of structural and functional imaging. Conventional and diffusion MRI are performed routinely for patients presenting symptoms of brain tumours (14). Conventional MRI provides information about anatomical structure, and therefore termed structural imaging. On the other hand, diffusion MRI probes the diffusion of molecules (mainly water) to reveal tissue microstructure and is therefore classified as a

functional imaging technique. Additionally, a variety of imaging sequences can be adjusted to emphasise different properties of tissues and exhibit complementary information from abnormal tissues. For example, primary brain tumours generally appear hyperintense on T2-weighted images while they appear hypointense on T1-weighted images (15). In terms of patient safety, MRI offers non-ionising radiation which is desirable in children as they are more susceptible to radiation side effects than adults (10, 16). As a result, MRI including conventional and diffusion MRI are considered in this thesis.

In conventional clinical diagnosis, radiologists examine imaging scans based on a set of known tumour characteristics and report the histological tumour type from these findings (17, 18). A review of recent radiological reports has revealed that the overall diagnostic classification accuracy for childhood brain tumours is 67%. In terms of individual tumour types: pilocytic astrocytomas, medulloblastomas and ependymomas achieved 75%, 73% and 25% respectively (19). These diagnostic accuracies depend on a combination of training, expertise and judgement of the radiologists, which introduce variation of clinical-decision making. However, the use of computer-aided analysis has been shown to improve upon radiological classification accuracy while improving processing time (18-20). Several studies have reported the usefulness of texture analysis (TA), a quantitative measurement of surface arrangement, for identification of pathological conditions such as multiple sclerosis, head injury and brain tumours (21-24).

Some studies using conventional MRI reported less distinctive information for certain tumour types and low discriminative power between high-grade and low-grade tumours

(12, 24-26). This can be caused by the limited capacity in monitoring functional changes when using conventional MRI (27). Diffusion MRI has been shown to provide added histological information and higher diagnostic classification accuracy compared to structural MRI (28, 29). Recently diffusion MRI based histogram analysis has been reported to produce high classification accuracy of paediatric posterior fossa tumours (30). Therefore it is of interest to examine the performance of a broader range of TA techniques using a data set containing tumours from other regions of the brain. It is also desirable to automate some of the manual steps in the presented processing pipeline.

Although there are several studies using either conventional or diffusion MRI based TA, the combination of textural featured derived from both imaging approaches has not been widely studied in the context of paediatric brain tumours. We hypothesise that given the complementary nature of the information contained in the output of the two imaging approaches, the integration of textural features derived from them may improve automated diagnosis.

1.2 Aim and Objectives

The issues discussed in the previous section raise the following research question: can the combination of textural features, derived from multimodal MRI, improve the classification of paediatric brain tumours? In attempting to answer this question, this thesis aims to develop and evaluate a framework for the texture analysis of multimodal MRI (e.g. conventional and diffusion MRI), in order to provide appropriate biomarkers to discriminate different tumour types for childhood brain tumours. To achieve this aim, three main objectives are defined as follows:

The first objective has two parts. The primary objective is to identify a suitable segmentation technique in order to enable the reproducibility of segmentation results, and speed up the segmentation process. The secondary objective is to identify which MR image type is the most appropriate for brain tumour segmentation based on the chosen method.

The second objective is to identify which MR-based image types can provide more effective TA-based classification of childhood brain tumours. Classification performance is statistically compared amongst seven types of MR images: T2, FLAIR, pre-contrast T1, post-contrast T1, ADC, MD and FA.

The third objective is to integrate the textural features from multi-centre, multimodal MRI, with the aim of improving the diagnostic classification accuracy of childhood brain tumours.

In order to answer the research question and accomplish the stated objectives, multidisciplinary topics are reviewed, researched and evaluated, as detailed in section 1.4. The fulfilment of the thesis objectives resulted in the original contributions described in the following section.

1.3 Thesis Contributions

The original contributions of this thesis enhance the diagnostic classification performance of paediatric brain tumours based on textural features of multimodal MRI. Several image processing and classification techniques were selected based on previous studies and

specific properties that are well suited for this domain. Diagnostic classification accuracy was further improved through the development of novel hybrid methods.

The dataset used in this study are conventional and diffusion MRI scans obtained from the Children's Cancer and Leukaemia Group (CCLG) database, operated by the CCLG Functional Imaging Group. The main contributions of this study, according to the three main objectives defined in section 1.2, are heighted below.

The first contribution is the comparison of semi-automatic segmentation methods: gradient vector flow snake and multiscale normalised cut method, evaluated against manual segmentation. The methods were applied on pre-contrast T1 and FA images, where multiscale normalised cut method was found to provide better tumour outlining. This contribution is presented in (31), shown in Appendix E.

The second contribution is the automated processing pipeline for TA of childhood brain tumours based on multimodal MRI. A preliminary study of a combined multiscale normalised cut and gradient vector flow snake method was built into the automated processing framework. This contribution is presented in (32), shown in Appendix F.

The third contribution involves the investigation of feature selection methods to classify childhood brain tumours based on textural features of conventional and diffusion MR images. The choice of feature selection method is important because it can affect the classification performance. This contribution is presented in (33), shown in Appendix G.

The fourth contribution demonstrates combining multi-centre conventional and diffusion MRI based texture for the characterisation of childhood brain tumours. The brain tumour

classification based on textural features of multimodal MRI was found to provide higher discrimination for ependymoma from medulloblastoma and pilocytic astrocytoma than using textural features of single modality MR-based images. This contribution is presented in (34), shown in Appendix H.

1.4 Thesis Organisation

The thesis is composed of nine chapters, with the introduction and conclusion in chapters one and nine, respectively. The literature review involving imaging principle and paediatric brain tumour characteristics based on conventional and diffusion MRI is included in chapter two. The theoretical background and methodology of processing methods used in this thesis are described in chapters three and four. The experimental work describing specific methods, results and discussion as well as conclusions are detailed in chapters five, six, seven and eight.

Chapter two presents a review of image acquisition and tumour characteristics based on conventional and diffusion MRI. The principles of these imaging techniques including the reconstruction of diffusion weighted imaging (DWI) and diffusion tensor imaging (DTI)-derived parameters are explained. Important imaging parameters are discussed, in order to understand the effects of different MRI sequences and associated imaging features of brain tumours. The characteristics of the most common childhood brain tumours: pilocytic astrocytomas, medulloblastomas and ependymomas on conventional and diffusion MRI are reviewed. The salient information and assumptions about these characteristics are discussed and illustrated.

Chapter three presents a literature review related to the processing framework in order to select appropriate methods for each step in the pipeline. The theoretical background of selected methods is described to provide an understanding of the methodology of the thesis.

Chapter four presents the thesis methodology, separating into four main experiments: pre-processing analysis, brain tumour segmentation, TA of single modality MR technique and TA of multimodal MRI. The quality control of the MRI dataset is also included.

Chapter five focuses on the pre-processing pipeline, which is primarily based on the FMRIB Software Library (FSL) framework. The process is composed of diffusion MRI-related artefact removal, skull stripping, image registration and intensity normalisation methods. An experiment is designed to find the optimal parameters and techniques to fit into the pre-processing framework.

Chapter six introduces a novel hybrid semi-automatic segmentation method. This technique is based on the combination of graph-cut and active contour based segmentation. The proposed method is evaluated by two approaches. The first approach involves a scoring system where four trained observers rank the effectiveness of the segmentation. The second is a supervised evaluation approach measuring similarity between manual segmentation and semi-automatic segmentation.

Chapter seven presents an experiment conducted to evaluate the effectiveness of brain tumour classification using textual features obtained from individual MR image types (single modality). Two feature selection methods; principal component analysis and the combined mutual information-based feature selection and sequential forward selection

method are assessed to find the optimal feature selection technique for the dataset. A support vector machine is used as the classifier. The comparison between the two feature selection methods and the classification performance of individual MR image-based TA are examined by a hypothesis test for two proportions. The performance of each MR image type is also evaluated, using the misclassification error rate generate by a well-known validation technique.

Chapter eight presents the integration of multimodal MRI based textural features for the classification of brain tumours. A single modality and multimodal MRI based texture classification is compared using classification outcomes and hypothesis test for two proportions. In terms of the classifier a multinomial logistic regression method is employed to build logistic models for the recommend processing frameworks. The images of misclassified cases are exhibited to demonstrate difficulties of the classification approaches.

1.5 Summary

This chapter presented the motivation of the research, the choice of imaging modality and a framework for the computer-aided diagnostic classification of childhood brain tumours. The research question and resolution approaches were described in the aim and objectives section. The final section listed the thesis contributions and organisation, in order to guide the reader through the content presented in later chapters.

CHAPTER 2 IMAGE ACQUISTION AND CHARACTERISTICS OF PAEDIATRIC BRAIN TUMOURS

This chapter presents image acquisition principles and characteristics of paediatric brain tumours based on conventional and diffusion MRI in order to support the three main objectives of this thesis. A number of important parameters for MRI acquisition are discussed, providing insight into imaging characteristics of the three most common childhood brain tumours. Learning the MRI features of these brain tumours provide a deeper understanding of the meaning of the image textural features, extracted from individual histological tumour types.

2.1 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a modern clinical imaging technique, which exploits the fundamental principles of Nuclear Magnetic Resonance (NMR) phenomenon. NMR was investigated since early 20th century by various groups of scientists. Felix Bloch and Edward Mills Purcell independently developed breakthrough methods for precise measurements of nuclear magnetic resonance frequency and magnetism of atomic nuclei, which was used to study the composition of different materials (35, 36). This piece of work was awarded the joint Nobel Prize in Physics in 1952 to both of them (37). Richard Robert Ernst primarily investigated nuclear magnetic resonance spectroscopy to study chemical structure of substances and was awarded the Nobel Prize in Chemistry in 1991 (38). Later, Raymond Damadian invented the first MR scanner and demonstrated

the application of NMR in living systems (39). Subsequently, Paul C. Lauterbur and Sir Peter Mansfield employed MRI for medical diagnostics and research, which produced a tremendous impact on the medical research field. With their findings, they were granted the shared Nobel Prize in Physiology or Medicine in 2003 (40).

MR-based image acquisition and reconstruction depends on physical device components, processes and associated parameters within the context of a MR experiment. Four important components of a MRI scanner: the static magnetic field of the scanner, the radio frequency pulse, the gradients, and the computer system, are discussed as graphically shown in Figure 2.1. The following sections discuss these fundamentals in detail, adapted from (41-46).

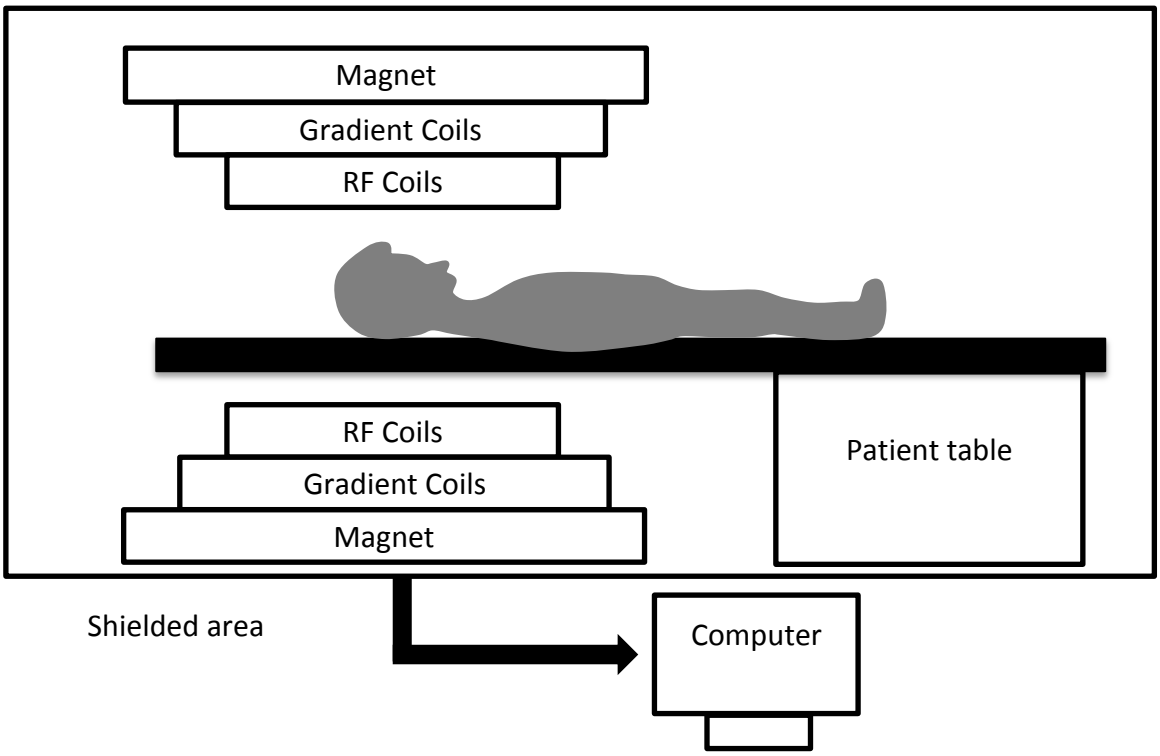


Figure 2.1: A simplified diagram of a MRI system. It composes of four main components: magnet, gradient coils, RF coils and computer system. Imaging is done in a shielded area (Adapted from (41)).

2.1.1 Nuclei in a Magnetic Field

A nucleus with an odd atomic number (odd number of protons) is a charged particle. It creates a magnetic field and rotates about its own axis (Figure 2.2). The rotation of a nucleus is called nuclear spin or spin which produces a magnetic moment (nuclear magnetic dipole moment) ($\boldsymbol{\mu}$) and an angular momentum ($\boldsymbol{p}$). The magnetic moment and angular momentum are linearly related with a factor termed the gyromagnetic ratio (γ), as shown in equation (2.1). The gyromagnetic ratio depends on the rotational frequency of an individual nucleus, placed in an external magnetic field.

$$\boldsymbol{\mu} = \gamma \boldsymbol{p} \tag{2.1}$$

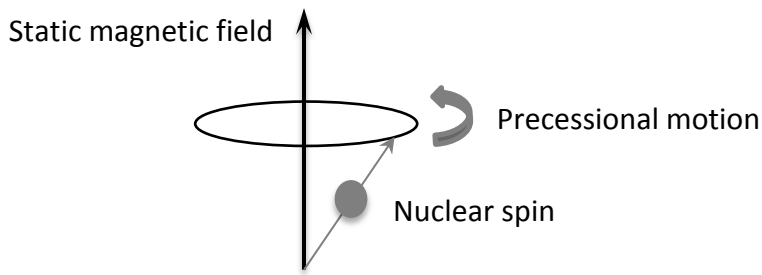


Figure 2.2: Precession of a nucleus around its own static magnetic field axis.

When an external magnetic field $\widehat{\boldsymbol{B}}_{\boldsymbol{0}}$ is exerted, the magnetic moment $\boldsymbol{\mu}$ experiences a torque (turning force) $\widehat{\boldsymbol{F}}$, forcing the spins to precess around the $\widehat{\boldsymbol{B}}_{\boldsymbol{0}}$ axis. The torque $\widehat{\boldsymbol{F}}$ is the cross product of the magnetic moment and magnetic field given in equation (2.2).

$$\widehat{\boldsymbol{F}} = \boldsymbol{\mu} \times \widehat{\boldsymbol{B}}_{\boldsymbol{0}} \tag{2.2}$$

The torque alters the angular moment with the rate of change given by equation (2.3).

$$\frac{d\boldsymbol{p}}{dt} = \boldsymbol{\mu} \times \widehat{\boldsymbol{B}}_{\boldsymbol{0}} \tag{2.3}$$

From equation (2.1), $\boldsymbol{p} = \boldsymbol{\mu}/\gamma$, provides the rate of change in the magnetic moment about the $\widehat{\boldsymbol{B}}_{\boldsymbol{0}}$ axis, derived in equation (2.4).

$$\frac{d\boldsymbol{\mu}}{dt} = \boldsymbol{\mu} \times \gamma \widehat{\boldsymbol{B}}_{\mathbf{0}} = \boldsymbol{\mu} \times \boldsymbol{\omega}_{\boldsymbol{L}} \tag{2.4}$$

where $\boldsymbol{\omega}_{\boldsymbol{L}}$ is called the Larmor frequency.

The Larmor frequency is the natural resonance frequency of a spin system and is unique for each nucleus; for example ^{1}H and ^{13}C resonate at 42.58 and 10.71 MHz at 1 Tesla (T) of $\widehat{\boldsymbol{B}}_{\mathbf{0}}$ respectively (42). However, nuclei with the same γ will have different Larmor frequencies at different spatial locations, when $\widehat{\boldsymbol{B}}_{\mathbf{0}}$ is inhomogeneous, as explained later in gradient encoding.

The spins can rotate in parallel with lower energy (shorter, thinner vector) or anti-parallel with higher energy (longer, thicker vector) around the $\widehat{\boldsymbol{B}}_{\mathbf{0}}$ axis (Figure 2.3 (a)). More number of nuclei precess with lower energy, as shown in Figure 2.3 (b). The vector sum of all magnetic moments (both parallel and anti-parallel) is the net magnetisation $(\widehat{\boldsymbol{M}})$, which is parallel to the main magnetic field and aligns with the z axis as convention (equation 2.5).

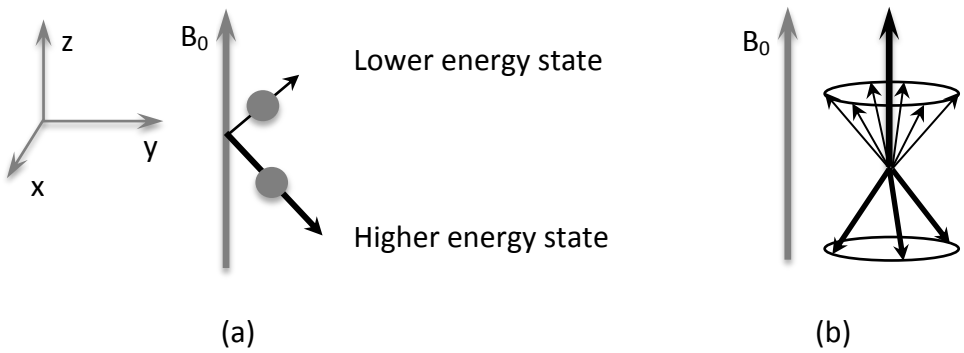


Figure 2.3: Energy states and magnetisation of nuclear spins. When nuclei precess around magnetic field, (a) protons align parallel with B_0 and anti-parallel with B_0, (b) the net magnetisation is parallel to the main magnetic field. A longer and thicker vector indicates higher energy state.

$$\widehat{\boldsymbol{M}} = \sum_{i=1}^{N_s} \boldsymbol{\mu_i} \tag{2.5}$$

where $\boldsymbol{\mu_i}$ are the magnetic moments, $i \in \{1,2, \dots N_s\}$ and N_s is the total number of spins.

2.1.2 Radio Frequency Pulse Excitation and Pulse Sequences

When a Radio Frequency (RF) pulse is transmitted; the spin system, precessing at Larmor frequency, is perturbed. A 90° RF pulse flips the spins onto the transverse x-y plane (Figure 2.4 (a)) and causes the spins to rotate at different Larmor frequency (Figure 2.4 (b)). The net magnetisation becomes nonequilibrium state and is decomposed to two magnetisation components: longitudinal $\widehat{\boldsymbol{M}}_{\boldsymbol{z}}$ and transverse $\widehat{\boldsymbol{M}}_{\boldsymbol{xy}}$.

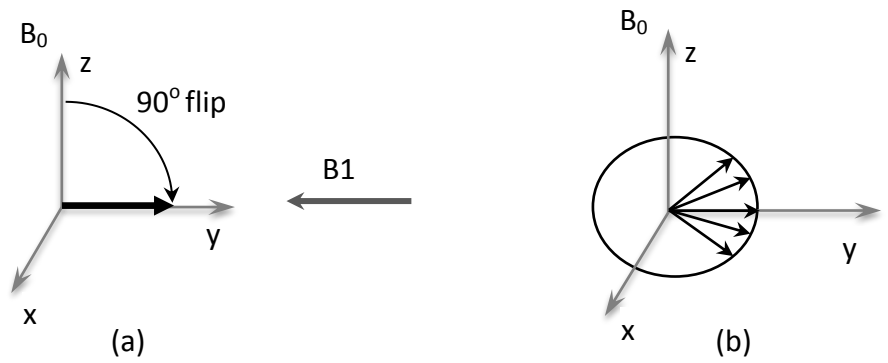


Figure 2.4: De-phasing of nuclear spins after a 90° RF pulse excitation. Nuclear spins are flipped with a 90° pulse at (a) time = t_0 and start de-phasing at (b) time > t_0.

After perturbation of the RF pulse, the spins return to the equilibrium state, by releasing excess energy into the surrounding environment. The rate of spins returning to the equilibrium state is different depending on the chemical composition of a measured tissue. The process of returning to the equilibrium is called longitudinal or spin-lattice relaxation where the transverse magnetisation component $\widehat{\boldsymbol{M}}_{\boldsymbol{xy}}$ decays to zero, and the

longitudinal magnetisation component $\widehat{\boldsymbol{M}}_{\boldsymbol{z}}$ returns to equilibrium magnetisation $\widehat{\boldsymbol{M}}_{\boldsymbol{0}}$. The dephasing of the spins in the transverse plane is called transverse or spin-spin relaxation.

The process of magnetisation $\widehat{\boldsymbol{M}}_{\boldsymbol{xy}}$ decaying to zero, which incorporates both spin-lattice and spin-spin relaxation, is exponential and called 'free induction decay' (FID). The value of $\widehat{\boldsymbol{M}}_{\boldsymbol{xy}}$ and $\widehat{\boldsymbol{M}}_{\boldsymbol{z}}$ can be derived after the application of an α degree RF pulse as defined in equation (2.6) and (2.7).

$$\widehat{\boldsymbol{M}}_{\boldsymbol{xy}}\,(t) = \widehat{\boldsymbol{M}}_{\boldsymbol{0}} sin\alpha e^{i(\omega_0 t+\varphi)} e^{-t/T2} \tag{2.6}$$

$$\widehat{\boldsymbol{M}}_{\boldsymbol{z}}(t) = \widehat{\boldsymbol{M}}_{\boldsymbol{0}}(1-(1-cos\alpha)e^{-\frac{t}{T1}}) \tag{2.7}$$

$T2$ is the transverse relaxation time (or spin-spin relaxation time) and indicates the time taken for the transverse magnetisation to decay to 36.7% of its initial value $\widehat{\boldsymbol{M}}_{\boldsymbol{0}}$. $T1$ is the longitudinal relaxation time (or spin-lattice relaxation time) and represents the time taken for the longitudinal magnetisation to recover to a value of 63.2% of the equilibrium value $\widehat{\boldsymbol{M}}_{\boldsymbol{0}}$. Transverse relaxation time is generally much faster than longitudinal relaxation. The values of $T1$ and $T2$ are considered intrinsic properties of tissues because they remain constant for a specific tissue at a given uniform magnetic field strength.

The rate of the FID can be described by another relaxation time, $T2^*$ (42), as shown in equation (2.8). The value of $T2^*$ is the relaxation time which varies depending on the spin-spin relaxation time and inhomogeneities in the main magnetic field. This value is an intrinsic property of each tissue type at the given magnetic field (47).

$$\frac{1}{T2^*} = \frac{1}{T2} + \frac{1}{T2'} \tag{2.8}$$

where $T2'$ is the time taken for the signal to decay due to the field inhomogeneity.

In MRI acquisition, multi-pulse experiments such as gradient-echo, spin-echo, and inversion-recovery sequences are commonly used. Gradient-echo is the simplest pulse sequence and provides faster image acquisition than other sequences (48). It consists of a series of α^o pulses, where each sequence is separated by repetition time (TR). The α^o is less or equal to 90° and typically ranges between 10° to 80° (49).

The spin-echo pulse sequence is one of the most commonly used pulse sequences for generating T1- and T2-weighted images. It starts with a 90° pulse and is followed by a 180° pulse after a period of TE/2 (TE is time of echo). After the 180° pulse, the spin-echo pulse sequence is repeated. A diagram of the spin-echo pulse sequence and its rotating magnetisation are shown in Figure 2.5.

The inversion-recovery pulse sequence starts with a 180° pulse inverting the longitudinal magnetisation against the main magnetic field and followed by a 90° pulse that brings the residual magnetisation into the transverse plane. The time between the initial 180° pulse and the following 90° pulse is called inversion time (TI). The duration of TI allows recovery of the magnetisation to equilibrium and also determines the signal attenuation in particular tissue. The inversion-recovery sequence suppresses fluids at TI ranging between 2000 to 2900 milliseconds depending heavily on TR and pathological condition (50). With this property, it is used to generate fluid attenuated inversion recovery (FLAIR) images in order to enhance contrast between lesion and fluid. A schematic of inversion recovery sequence and its rotating of magnetisation are shown in Figure 2.6.

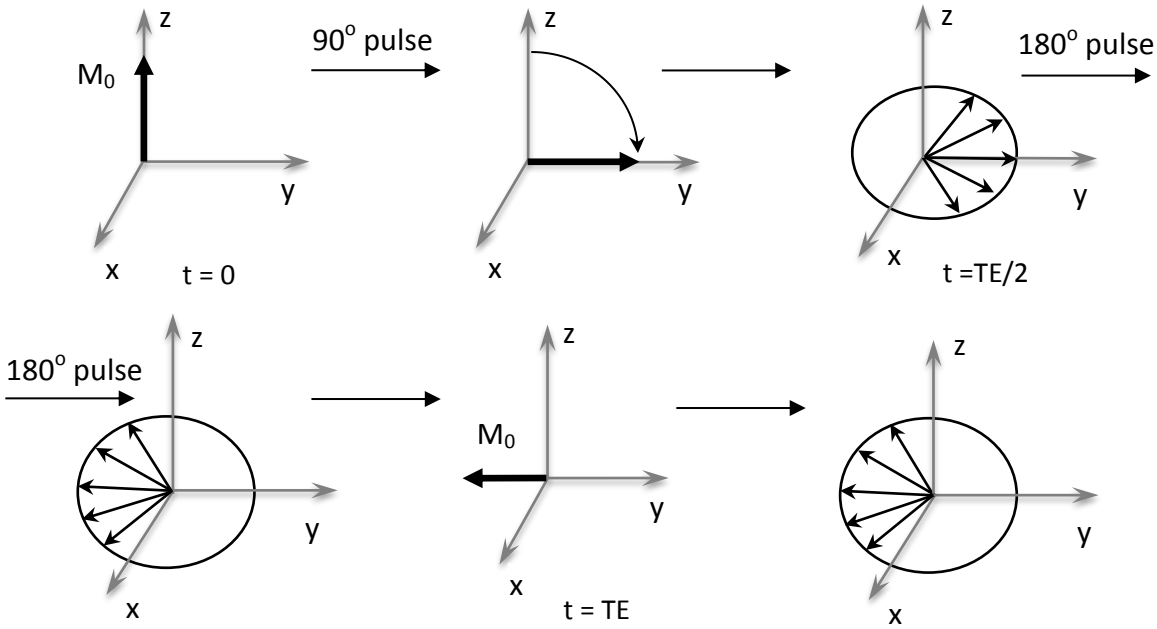


Figure 2.5: The change in magnetisation of a spin-echo sequence (reproduced from (42)). Nuclear spins are flipped with a 90° pulse and simultaneously de-phasing till time = TE/2. The nuclear spins are then flipped with a 180° pulse and rephased at time = TE.

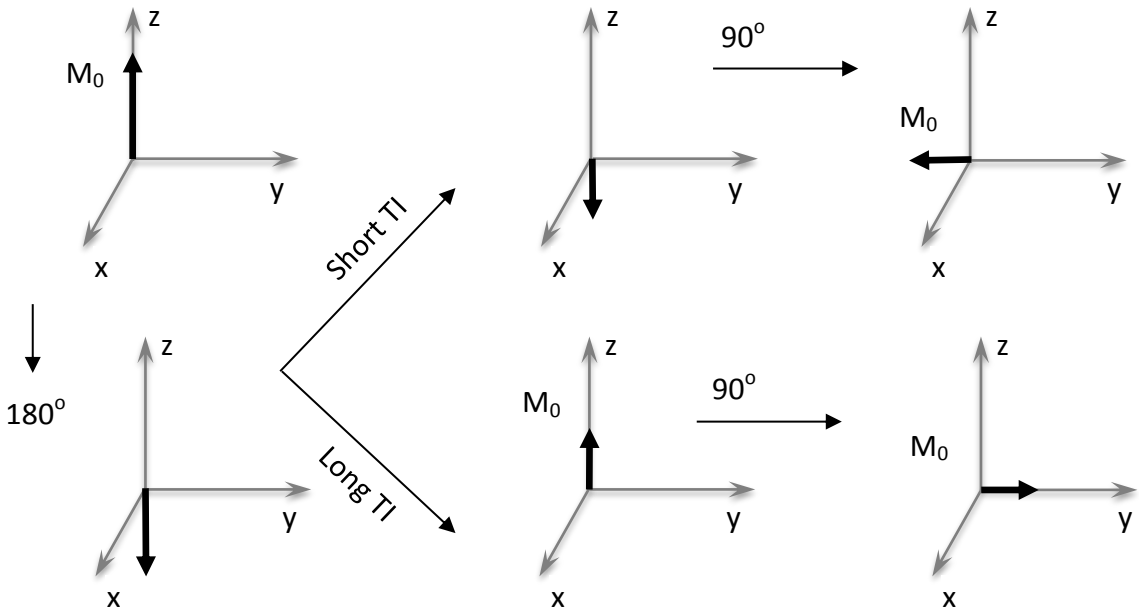


Figure 2.6: The change in magnetisation of an inversion recovery sequence (reproduced from (42)). The longitudinal magnetization is inverted by 180° and rephased onto the transverse plane after time = TI.

2.1.3 Gradient Encoding and Spatial Localisation

After transmitting the RF pulse, magnetic field gradient pulses are emitted to control the spatial information of MR signal. The gradient coils produce three main magnetic field gradient including frequency encoding or readout gradient ($\widehat{\boldsymbol{G}}_{\boldsymbol{x}}$), phase encoding gradient ($\widehat{\boldsymbol{G}}_{\boldsymbol{y}}$) and slice-selective gradient ($\widehat{\boldsymbol{G}}_{\boldsymbol{z}}$). The amplitudes of magnetisation $\boldsymbol{B}$ vary linearly with position along a chosen axis as shown in equation (2.9).

$$\boldsymbol{B} = \widehat{\boldsymbol{B}}_{\boldsymbol{0}} + x\widehat{\boldsymbol{G}}_{\boldsymbol{x}} + y\widehat{\boldsymbol{G}}_{\boldsymbol{y}} + z\widehat{\boldsymbol{G}}_{\boldsymbol{z}} \tag{2.9}$$

When the gradient pulse is applied, the spins precess at different Larmor frequencies at different spatial locations. Frequency encoding gradient, $\widehat{\boldsymbol{G}}_{\boldsymbol{x}}$ causes the change in resonance frequency along the x axis as shown in equation (2.10). It contributes to the real part of complex values in k-space, explained later in section 2.1.4.

$$\boldsymbol{\omega}(x) = \gamma(\widehat{\boldsymbol{B}}_{\boldsymbol{0}} + x\widehat{\boldsymbol{G}}_{\boldsymbol{x}}) \tag{2.10}$$

In order to encode information onto k-space, a phase encoding gradient, $\widehat{\boldsymbol{G}}_{\boldsymbol{y}}$ is applied to acquire imaginary part of k-space. When $\widehat{\boldsymbol{G}}_{\boldsymbol{y}}$ is switched on for a short period, t_y; this results in the spins having different phase angles (ϕ) at different y positions according to equation (2.11).

$$\phi(y) = -\gamma y\widehat{\boldsymbol{G}}_{\boldsymbol{y}}t_y \tag{2.11}$$

The slice-selective gradient, $\widehat{\boldsymbol{G}}_{\boldsymbol{z}}$ varies the resonant frequency along the z direction (equation (2.12)), which is used to control thickness and position of an image slice.

$$\boldsymbol{\omega}(z) = \gamma(\widehat{\boldsymbol{B}}_{\boldsymbol{0}} + z\widehat{\boldsymbol{G}}_{\boldsymbol{z}}) \tag{2.12}$$

The thickness and position of the image slice can be varied by adjusting the amplitude of $\widehat{\boldsymbol{G}}_{\boldsymbol{z}}$ and the bandwidth of the RF pulse, i.e. the larger the amplitude of $\widehat{\boldsymbol{G}}_{\boldsymbol{z}}$ or the narrower

the bandwidth of RF pulse ($\Delta\omega$), the thinner the slice. The slice thickness can be calculated as shown in equation (2.13).

$$\Delta z = \Delta\omega/\gamma|\widehat{\boldsymbol{G}}_{\boldsymbol{z}}| \tag{2.13}$$

So far, the radiofrequency and magnetic field gradient pulses were reviewed. In addition, a schematic diagram of theses pulses is drawn to illustrate their sequence timings of gradient echo sequence, spin echo sequence and inversion recovery sequence, as shown in Figure 2.7 to 2.9, respectively.

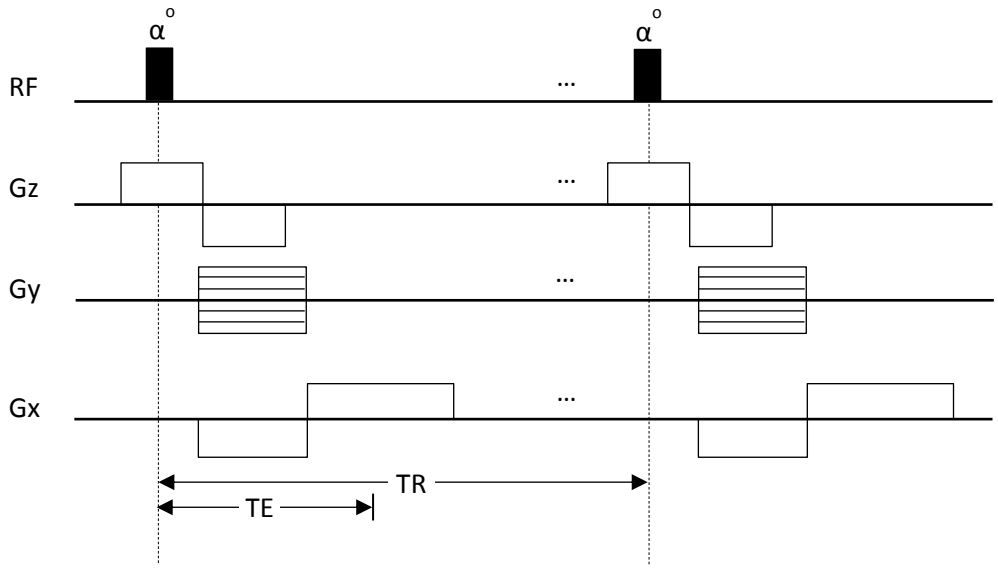


Figure 2.7: A typical gradient-echo imaging sequence. The sequence is composed of pulses having flip angle less than 90° (Adapted from (46)).

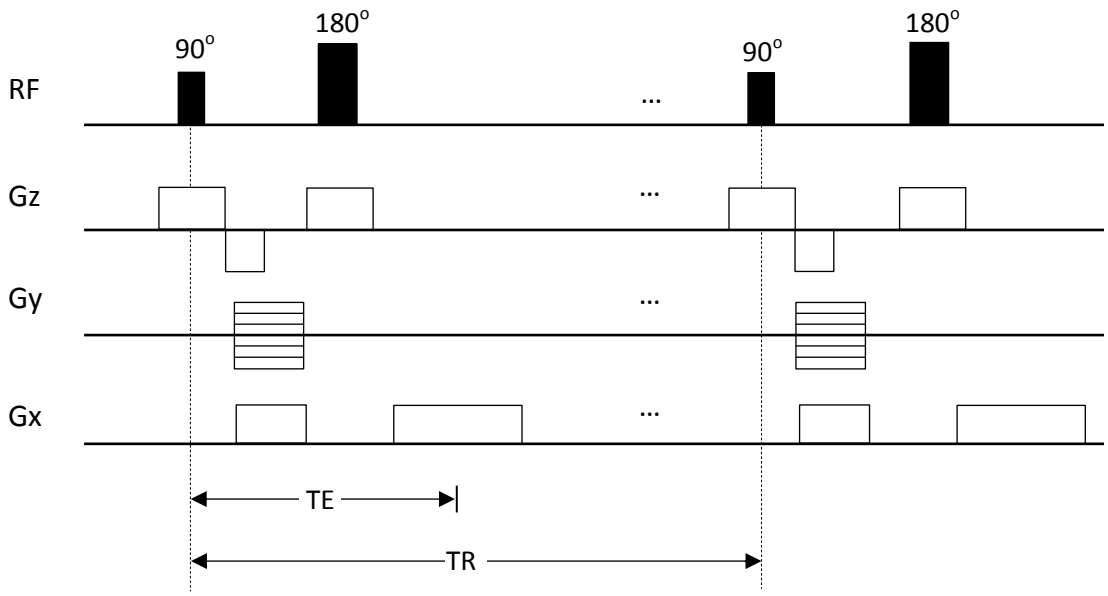


Figure 2.8: A typical spin-echo imaging sequence. The sequence starts with 90° pulse and followed by a 180° pulse after a period of TE/2 (Adapted from (46)).

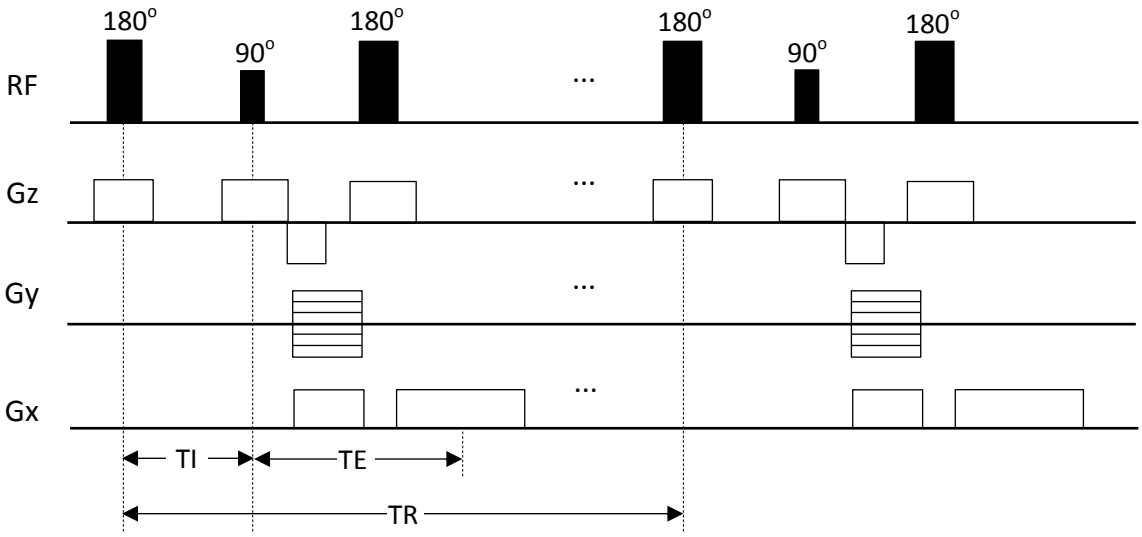


Figure 2.9: A typical inversion recovery imaging sequence. The sequence starts with a 180° pulse and followed by a 90° pulse after a period of TI (Adapted from (46)).

2.1.4 MR Image Reconstruction

The fourth component in MRI is the computer system, which is used to convert the analogue RF signal to digital signal through an analogue/digital convertor. After

converting analogue RF signal, the digital MR signal is stored in a k-space domain containing a raw data space of complex values. Subsequently, the image space can be constructed from the k-space by a Fourier transform, as shown in Figure 2.10.

Given a k-space with a matrix of size N x M, an element in k-space is composed of a real (k_x) and an imaginary (k_y) value according to the equations (2.14) and (2.15), respectively.

$$k_x = \gamma N \Delta t G_x \tag{2.14}$$

$$k_y = \gamma M \tau G_y \tag{2.15}$$

where γ is the gyromagnetic ratio, N is the number of samples along the read gradient, M is the number of times the phase encoding transmitted, Δt is the sampling time interval, G_x is frequency encoding, G_y is phase encoding, τ is the duration of G_y.

In the clinical setting, anatomical images are obtained from three different views: transverse (axial), coronal and sagittal. Axial view is more commonly acquired for both structural and functional imaging, as recorded in the CCLG database. Only axial imaging is used in this thesis.

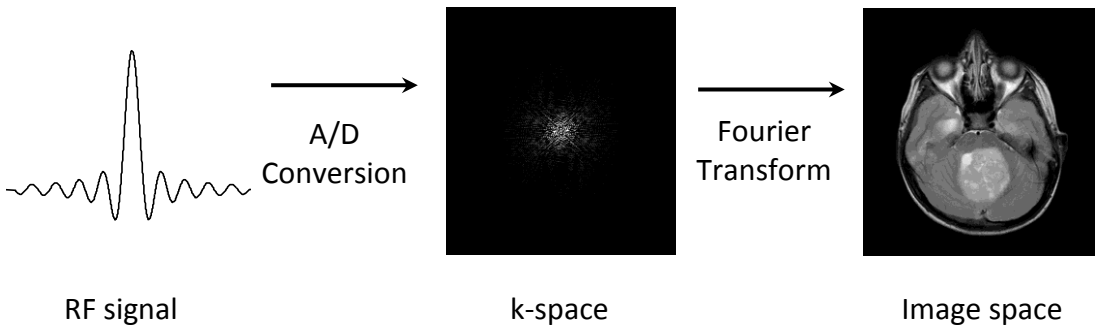


Figure 2.10: A conversion of RF signal to a digital image. An analogue RF signal is digitized and recorded in a k-space domain before performing Fourier transform to obtain the image data (case MB532).

2.1.5 Image Contrast

MRI scans derived from different pulse sequences have distinctive image contrast. The three common pulse sequences are the gradient-echo, spin-echo and inversion recovery, producing particular signal intensity according to equations (2.16), (2.17) and (2.18) (51, 52). T1- and T2-weighted images are generally scanned with a spin-echo sequence and FLAIR is acquired using an inversion recovery technique.

In a gradient-echo technique

$$I \propto N(H)(e^{-\frac{TE}{T2^*}})(1 - e^{-\frac{TR}{T1}}) \tag{2.16}$$

In a spin-echo technique

$$I \propto N(H)(e^{-\frac{TE}{T2}})(1 - e^{-\frac{TR}{T1}}) \tag{2.17}$$

In inversion recovery technique

$$I \propto N(H)(1 - 2e^{-\frac{TI}{T1}} + e^{-\frac{TR}{T1}}) \tag{2.18}$$

where $N(H)$ is the proton density, TR is the repetition time of a pulse sequence and TE is the period between RF pulse and spin echo, $T1$ is the longitudinal relaxation time, $T2$ is traverse relaxation time, $T2^*$ is relaxation time due to spin dephasing and inhomogeneity of magnetic field, and TI is the inversion time.

TR, TE, TI are defined by the pulse sequence design and can be adjusted to control T1, T2 weighting and therefore tissue contrast. A T1-weighted image is generally generated by using short TR to pronounce T1-weighted contrast and short TE to exclude T2-weighting effects. A T2-weighted image is excited by using long TR to remove T1-weighting and long

TE to introduce T2-weighted contrast (12, 53). A FLAIR image is produced by using long TI to null the water signal and long TR to eliminate T1 effects.

The T1 and T2 relaxation times have distinct duration in different tissue types. The longest to shortest period are in fluid, water-based tissue and fat-based tissue respectively. However, T2 is generally shorter than T1 and shows inverted signal intensity for a given tissue (41). The lengthened T1 values tend to decrease signal intensity, while the lengthened T2 values tend to increase signal intensity. As a result, grey-level intensity of brain components is visually inverted between T1- and T2-weighted, for example CSF appears bright on the T2-weighted image (Figure 2.11(a)) and dark on the T1-weighted image (Figure 2.11(b)). FLAIR is a T2 sequence with suppressed fluid signal appearing as low intensity (Figure 2.11(c)). The further imaging characteristics of brain tumours based on MRI is reviewed in section 2.3.

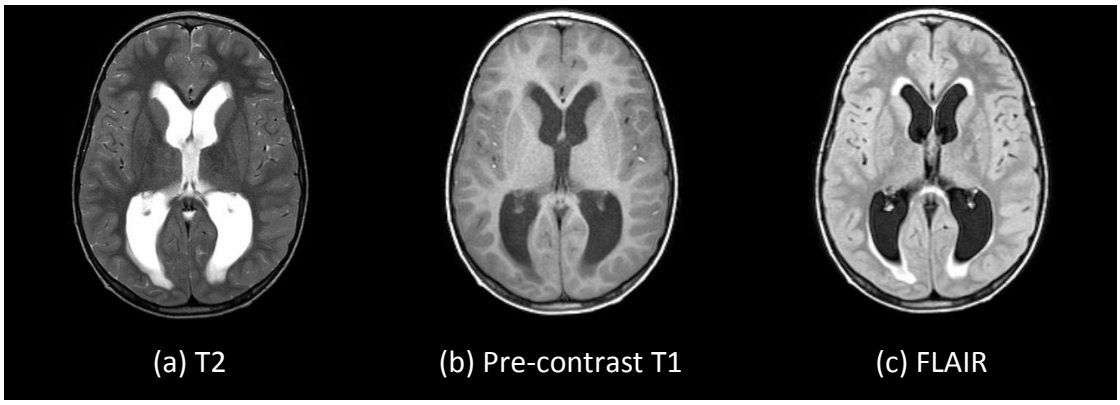


Figure 2.11: Complementary image contrast of brain tissue acquired based on T2, pre-contrast T1 and FLAIR (case MB718).

2.2 Diffusion Magnetic Resonance Imaging

Diffusion MRI is a non-invasive functional imaging technique, which has been employed to investigate the tissue microstructure reflected by diffusion of water molecules. The following sections review the basic principle of the diffusion phenomenon and discuss the reconstruction of diffusion MRI.

2.2.1 The Basics of Diffusion

Diffusion is an essential process in all living organisms that require the transport of metabolites and water into cells. Studying diffusion has the potential to provide further insight into both cell physiology and tissue structures. In unhindered space, the diffusion of randomly mobile molecules is described by Brownian motion. Diffusion starts from a volume with a higher concentration of substances and is directed to a volume of lower concentration until equilibrium is established. The rate of change of this concentration is explained by Fick's first law (54) which states that the flux density is linearly proportional to the concentration gradient as shown in equation (2.19).

$$J = -D\frac{\partial \emptyset}{\partial x} \tag{2.19}$$

where J is the diffusion flux (amount of substance per unit area per unit time, e.g. mol/mm^2/s), D is the diffusion coefficient or diffusivity (e.g. mm^2/s), $\emptyset$ is the concentration of substance (e.g. mol/mm^3) and x is the displacement of the substance (e.g. mm).

As the diffusion of molecules is randomly distributed in unrestricted space, the probability profiles of molecular displacement follow a bell-shaped or Gaussian form. The probability

displacement is related to the diffusion time which was formularised by Albert Einstein in 1905 (55, 56) as shown in equation (2.20).

$$\langle r^2 \rangle = 2Dt \tag{2.20}$$

where $\langle r^2 \rangle$ is the mean-squared displacement, D is the diffusion coefficient, t is the observation time.

2.2.2 Diffusion Weighted Imaging

Diffusion in a living system is an intrinsic property of tissues that is independent of magnetic resonance relaxation properties (e.g. longitudinal and transverse relaxation) (57, 58). The effects of diffusion on the attenuation of precessing nuclear spins were firstly reported by Erwin Hahn in 1950 (59). The precessing nuclear moments contained in liquid molecules, mostly of low viscosity, are not only attenuated by the influence of longitudinal time and transverse time but also de-phased by the self-diffusion of the molecules when placed in an inhomogeneous magnetic field (58).

To obtain information about molecular diffusion in an imaging experiment, a diffusion-encoding gradient is applied as an additional readout gradient. The addition of the diffusion gradient leads to compression of the MR signals in the time domain and produces a poorer signal-to-noise ratio (SNR) (58). The common imaging technique used in diffusion MRI is Echo Planar Imaging (EPI) which can be the application of multiple sequences of gradient-echo or spin-echo sequences.

For a fixed diffusion weighting and a single diffusivity, it can be shown that the signal in a diffusion-weighted experiment is given by equation (2.21).

$$I = I_0 e^{-TE/T2} e^{-bD} \tag{2.21}$$

where $T2$ is the transverse relaxation time, I_0 is the signal intensity in the absence of any $T2$ or diffusion weighting, TE is the echo time and b is the 'b-factor' or 'b-value'.

The approximation of the diffusivity of DW-MRI is independent of the T2. This diffusivity, D is quantitatively derived as follows.

$$I_1 = I_0 e^{-TE/T2} e^{-b_1 D} \tag{2.22}$$

$$I_2 = I_0 e^{-TE/T2} e^{-b_2 D} \tag{2.23}$$

$$D = \frac{-1}{(b_2 - b_1)} \ln\left(\frac{I_2}{I_1}\right) \tag{2.24}$$

where I_1 and I_2 are signals in a diffusion weighted experiment at b-values of b_1 and b_2 respectively.

2.2.2.1 Derivation of Apparent Diffusion Coefficient

During the DWI procedure, a target image frame is divided into small and equal imaging voxels. In each imaging voxel, a diffusion-weighted intensity varies according to the rate of attenuation of precessing nuclear spins. Spin attenuation varies depending on tissue properties, for example diffusion attenuation in CSF is much greater than that observed in white matter because the diffusion of water molecules is relatively unhindered in the CSF than that in the white matter.

After DWI acquisition, a reconstruction map namely Apparent Diffusion Coefficient (ADC) can be calculated by using DWI at two b-values as shown in the equation (2.24). An ADC map measures the magnitude of diffusion of molecules in mm^2/s and can be rewritten as equation (2.25).

$$ADC = -\left(\ln \frac{I_1}{I_0}\right) / (b_1 - b_0) \tag{2.25}$$

where I_1 and I_0 is the signal intensity with gradient having b_1 s/mm^2 and without the diffusion weighting having b_0 = 0 s/mm^2 respectively.

Both DWI and ADC maps provide measurement of molecular diffusion. However, a DWI and an ADC image exhibit dissimilar image intensity on restricted and unrestricted space as shown in Figure 2.12.

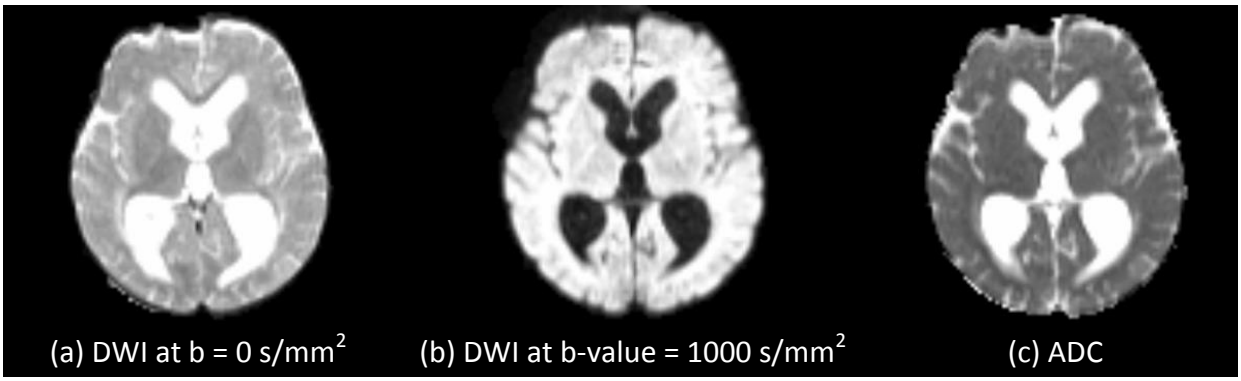


Figure 2.12: Signal intensity of restricted and unrestricted space based on DWI and ADC. An unrestricted area like CSF is bright on DWI at b = 0 s/mm^2 and an ADC map, but dark on DWI at b-value = 1000 s/mm^2 (case EP904). It is vice versa for a restricted area.

On areas with restricted diffusion, DWI appears bright, whereas ADC appears dark because accumulated nuclear spins in the area results in high signal on DWI and limited diffusion activity gives low signal intensity on ADC. On areas with unrestricted diffusion, DWI appears dark, whereas ADC appears bright because a small number of nuclear spins in the area result in low signal on DWI and unhindered diffusion gives high signal on ADC. The comparison of signal intensity obtained from DWI and ADC in restricted and unrestricted areas is tabulated in Table 2.1.

Table 2.1: Comparison of signal intensity of DWI and ADC images.

Image	Signal intensity in restricted area	Signal intensity in unrestricted area
DWI	Bright (Spins are more accumulated and yield higher signal.)	Dark (Less spins are caught and yield lower signal.)
ADC	Dark (Molecules cannot diffuse randomly; give low signal.)	Bright (Molecules can move more freely; give high signal.)

2.2.2.2 The Optimal b-Value

The b-value is a factor in diffusion weighted sequences and summarises the influence of the gradients on a DWI. The b-value can be estimated from the Stejskal-Tanner equation (2.26). The b-value is increased by either increasing the gradient strength (G) or the temporal separation of the gradients (δ).

$$b = (\gamma^2 G^2 \delta^2)\left(\Delta - \frac{\delta}{3}\right) \tag{2.26}$$

where γ is the gyromagnetic ratio and Δ is the time interval between the leading edges of the gradient lobes.

The higher the b-value, the stronger the diffusion weighting, which may show lesions more vividly but offers poorer SNR, as shown in Figure 2.13 and 2.14 because of longer TE, increased T2 and increased susceptibility to magnetic field gradient (60, 61). Susceptibility to magnetic field gradient accelerates the dephasing between protons, resulting in signal decay through T2* or severe image distortion. The optimal b-value was suggested by Bito et al. (62) to be 1.1/ADC which is considered as a rule of thumb. However, the gradient power is restricted as shown in the Stejskal-Tanner equation (2.26). To obtain the optimal b-value, the diffusion-encoding gradient must be increased, but this will result in poorer SNR as mentioned above. The optimal b-value in practice is

recommended to be in the range of 900 to 1200 s/mm^2 (61). In the clinical setting, the acquisition of diffusion MRI is commonly done at b = 0 as a reference, and up to 1000 s/mm^2 (generally found at 1000, 800 and 500 s/mm^2). A b-value of zero corresponds to an EPI T2-weighted image.

Generally an ADC map is acquired based on two b-values, which can be less accurate to reflect the pathology information than an ADC map derived from three b-values (e.g. b=0, b=500 and b=1000 s/mm^2) because the lower SNR of b=1000 s/mm^2 images presenting a higher standard deviation can be partially compensated by the median value of b=500 s/mm^2 (60).

Figure 2.13: The difference of imaging features and deterioration of SNR-based DWI at b-value = 0 and 800 s/mm^2 (case MB719).

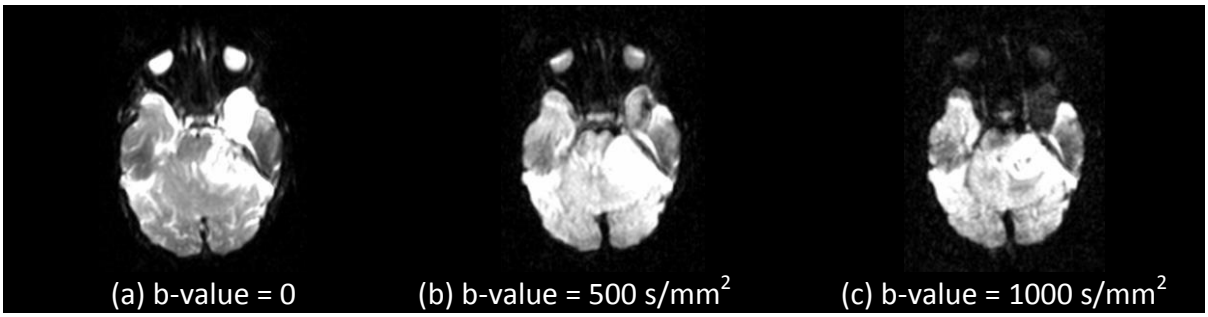


Figure 2.14: The difference of imaging features and deterioration of SNR-based DWI at b-value = 0, 500 and 1000 s/mm^2 (case MB950).

2.2.2.3 Rotational Variance

ADC in human white matter appears to depend on the direction of the applied diffusion-encoding gradient (63, 64). In the same region, if a subject moves when a particular diffusion-encoding gradient is applied, ADC values will be altered. This directional dependence or rotational variance of ADC could lead to complicated interpretation of the actual diffusion tissue in the particular region and could result in incorrect analysis. This issue can be addressed by using the rotationally invariant model of diffusion tensor imaging as described in the following section.

2.2.3 Diffusion Tensor Imaging

Diffusion Tensor Imaging (DTI) measures relative diffusion coefficients of water molecules in each pixel by scanning with DWI in at least six noncollinear directions (65). Compared with MRI, image contrast of DTI is more sensitive to fibre orientations and provides more information about fibre structure, particular in white matter.

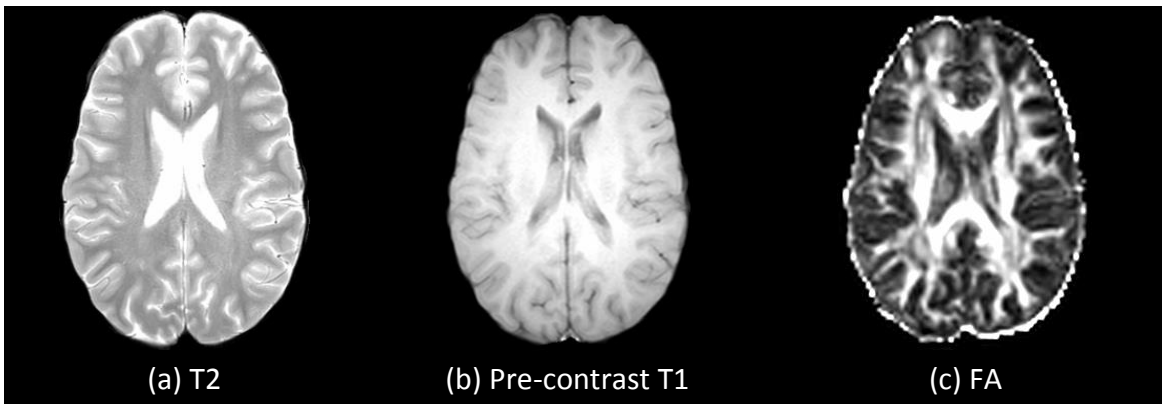


Figure 2.15: Comparison of white matter on T2, pre-contrast T1 and FA images. Images were acquired using a 1.5T MRI scanner (case MB622).

White matter on a T1- and a T2-weighted image appears relatively homogenous because of similar chemical composition within the tissue (42, 44), whereas a Fractional Anisotropy (FA) map, reconstructed from DTI shows more vivid patterns of white matter structure as shown in Figure 2.15. The sensitivity to subtle disturbance of white-matter tracts of DTI (66) could possibly provide information on tumour classification.

2.2.3.1 Diffusion Tensor Model

As mentioned in section 2.2.2, the ADC value has directional dependence. This can be compensated by using a diffusion tensor model of DTI which measures the principal diffusivity of molecules regardless of positions of fibres in space. The basic element of DTI used for such purposes is the diffusion tensor matrix ($\boldsymbol{D_m}$), a symmetric 3D matrix given by equation (2.27):

$$\boldsymbol{D_m} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix} \tag{2.27}$$

The diagonal elements correspond to the diffusivities along the three orthogonal axes (x, y, z) and the off-diagonal elements relate diffusivities along those.

The diffusion tensor depends on the weighting matrix (67) according to the relationship in equation (2.28).

$$ln\left(\frac{I_b}{I_0}\right) = -\sum_{i=1}^{3}\sum_{j=1}^{3} b_{ij}D_{ij} \tag{2.28}$$

where I_b is the signal intensity at b s/mm^2 of b-value, I_0 is the signal intensity at no diffusion gradient, D_{ij} are elements of the diffusion tensor matrix, $\boldsymbol{D_m}$ and the b_{ij} are elements of the weighting matrix (b-matrix), given by equation (2.29).

$$b_{ij} = \gamma^2 G_i G_j \delta^2 \left(\Delta t - \frac{\delta}{3}\right) \tag{2.29}$$

where G_i is an element of gradient and G_j is a transpose of G_i, γ is the gyromagnetic ratio, Δt is the time interval between the leading edges of the gradient lobes and δ is the temporal separation of the gradients.

The b-matrix is generated from a net effect of the gradient on each echo for each desired direction. The shape of gradient pulse can affect the b-matrix and diffusion images. Moreover, Leemans and Jones observed that the rotation of b-matrix can correct artefacts due to subject motion (68). However, the correction of the b-matrix is beyond the scope of this thesis.

2.2.3.2 Parameters Derived From the Diffusion Tensor Model

Reconstruction of the DTI maps requires at least six gradient directions to form a diffusion tensor matrix. A higher number of gradient directions, for example between 20 to 30, is recommended to estimate diffusion robustly (68). The diffusion tensor model can be graphically visualised as an ellipsoidal model with direction and length being annotated by eigenvectors and eigenvalues (λ). Three main eigenvalues: λ_1, λ_2 and λ_3, measure diffusivity along each of the three primary axes of the diffusion tensor matrix and are orthogonal axes of the ellipsoid. Fractional Anisotropy, FA and Mean Diffusivity, MD can be calculated from these eigenvalues, as follows:

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \tag{2.30}$$

$$FA = \sqrt{\frac{1}{2}}.\frac{\sqrt{(\lambda_1 - \hat{\lambda})^2 + (\lambda_2 - \hat{\lambda})^2 + (\lambda_3 - \hat{\lambda})^2}}{\sqrt{{\lambda_1}^2 + {\lambda_2}^2 + {\lambda_3}^2}} \quad (2.31)$$

where λ_1, λ_2 and λ_3 are eigenvalues of diffusion tensor, $\hat{\lambda}$ =$(\lambda_1 + \lambda_2 + \lambda_3)/3$.

2.2.3.3 Isotropy and Anisotropy

In each imaging voxel, the diffusion tensor model represents an ellipsoid in relation to local tissue structures. In some parts of the brain, the ellipsoid of the diffusion tensor has the same magnitude in all three main directions. Such a model has a spherical shape, representing isotropic diffusion. In other words, diffusion that is equally restricted or unrestricted in all directions results in low FA value. However, in some regions, the restriction of diffusion is highly directionally dependent. This scenario yields a high FA value, which is known as anisotropic diffusion.

In white matter, high anisotropy is frequently read because of uneven diffusivity with fast diffusivity along the fibre direction and slow diffusivity perpendicular to it. This effect is possibly influenced by the myelin sheath, local susceptibility gradients, axonal cytoskeleton, fast-axonal transport, and intact cell membranes (69). In grey matter and CSF, the diffusion is closer to isotropic because densely packed neural cells in grey matter or a hollow space of ventricle cause the diffusivity to be more likely the same in all directions (61).

2.2.3.4 Imperfection of Diffusion Tensor Imaging

Although DTI is dominantly superior in characterising fibre orientation in white matter, two assumptions of DTI could lead to misinterpretation of actual fibre orientation in some

scenarios (70). The first assumption is that the diffusion in white matter follows a Gaussian distribution. The diffusion is restricted in white matter where two or more fibre systems cross and therefore violate the Gaussian distribution. The second assumption is that it is sufficient to use a single diffusion tensor to characterise each voxel. The DTI model averages information from different diffusion compartments and reduces the certainty of tracking in partial-value pixels.

This means that the DTI model cannot interpret the fibre architecture when there is multiple fibre orientation within the same pixel. Low anisotropy can refer to either lack of coherent fibre orientation within the pixel or isotropic diffusion of water in unhindered space. In addition, when there is no water within a pixel, that pixel has no intensity. From the aforementioned assumptions, different types of neuroanatomy may lead to the same DTI result as graphically demonstrated in Figure 2.16.

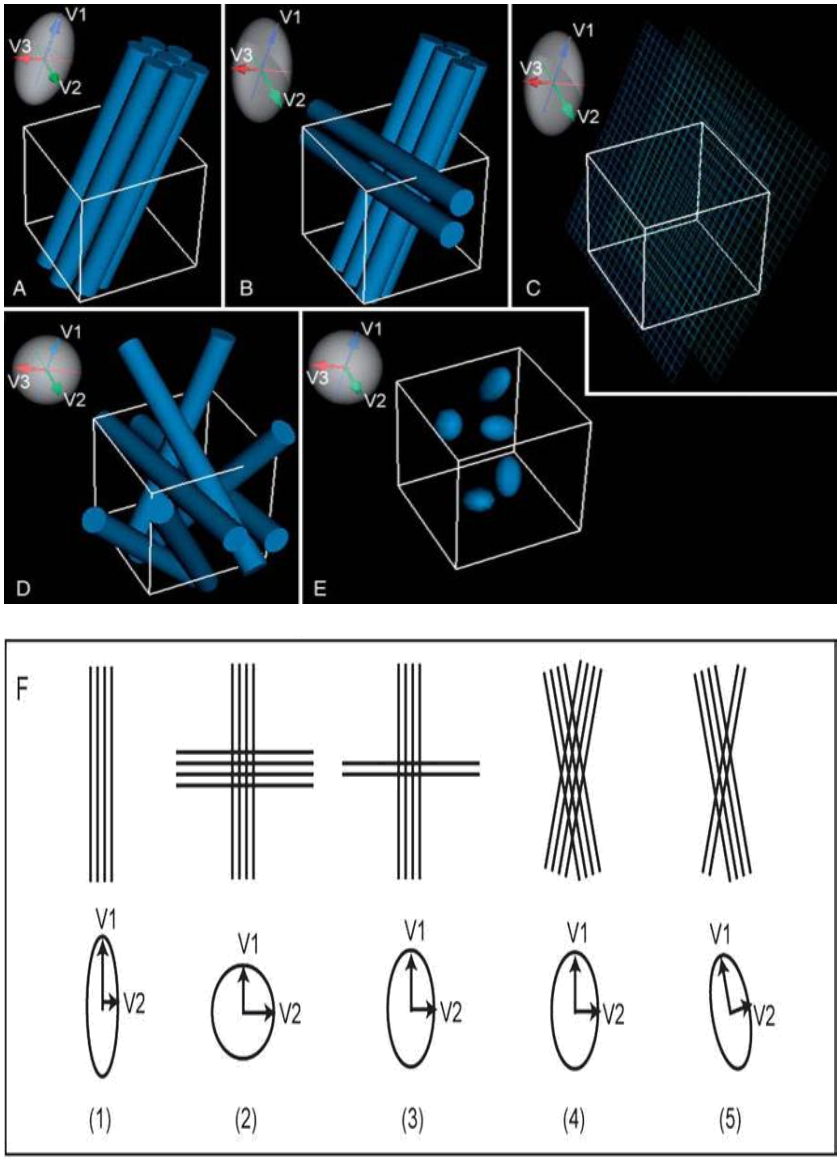


Figure 2.16: Fibre structures and shape of diffusion tensors (the blue cylinders denote fibre structures, the cubic frame denotes imaging voxel, the blue wireframe surface denotes planar shape of diffusion tensor). A: One fibre structure within an imaging voxel has a tubular shape. B: Two crossing fibre structures within an imaging voxel have less anisotropic diffusion. C: A planar shape of diffusion tensors has low permeability. D: Multiple fibre structures crossing with an imaging voxel produces spherical ellipsoid. E: The diffusion models have spherical shapes and no fibre structure. F: Arrangement of fibre structures represented by straight lines and their corresponding 2D diffusion ellipsoids with major and minor axis denoting the direction and magnitude of diffusion (Reproduced with permission from (44)).

2.3 Magnetic Resonance Imaging Characteristics of Paediatric Brian Tumours

2.3.1 Common Paediatric Brain Tumours

The most common paediatric brain tumours: pilocytic astrocytoma, medulloblastoma and ependymoma are considered. Pilocytic astrocytoma is a slow-growing tumour (grade I) and considered as the most common glioma, accounting for 20% of paediatric brain tumours (71). Medulloblastoma is a highly malignant and invasive embryonal tumour (grade IV) which represents 15-30% of all paediatric brain tumours (72). Ependymoma is the third most frequent histological type (grade I-III), following pilocytic astrocytoma and medulloblastoma. The incidence rate of ependymoma is 5-7% of all primary paediatric brain tumours (73).

In this section, the level of aggressiveness, cell of origin, anatomical location and imaging features of the brain tumours are described. In order to understand the biological aspects reviewed in the later sections, it is important to learn anatomical elements related to our study. Some important brain components are illustrated on a sagittal T2-weighted image as exhibited in Figure 2.17. The infratentorial and supratentorial region separate the lower and upper hemisphere of the brain. Tumours are most commonly found to be located infratentorially, in the posterior fossa[1]. However, some cases can migrate into the supratentorial brain region and spinal cord. Tumours which have different 'cells of origin', can originate from different sites, such as fourth ventricle and spinal cord.

[1] Posterior fossa is the region near the bottom of the skull.

Brain tumours are commonly composed of various components such as solid mass[2], cyst[3], calcification[4], oedema[5], necrosis[6] and haemorrhage[7]. All of these components are not necessarily present in every tumour. Only some components may be present with varying amounts from one to another. In visual analysis, the relative intensity of brain tumour components is simultaneously inspected on multimodal MRI. The relative intensity of brain tumours is described in relation to the intensity level of normal tissues of grey matter, white matter and cerebrospinal fluid as shown in Figure 2.17.

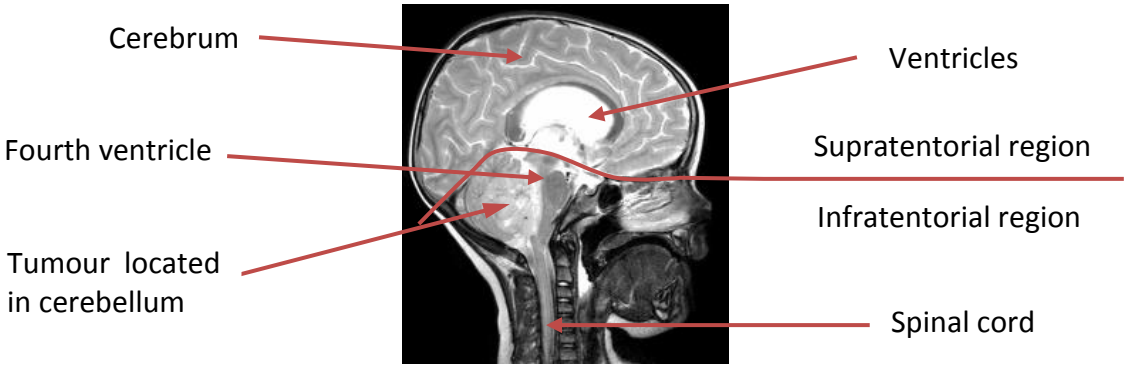


Figure 2.17: Some brain components and location on a sagittal T2-weighted image related to this thesis (case MB867).

Figure 2.18 shows grey-level intensity of grey matter, white matter and CSF on conventional and diffusion MRI. Grey matter is relatively brighter than white matter on T2 FLAIR, DWI at b-value = 0 and DWI at b-value = 1000 s/mm^2 , as shown in Figure 2.18 (a), (b), (g) and (h) respectively.

[2] Solid mass is an abnormal tissue that does not contain liquid.
[3] Cyst is a closed sac tissue that may be filled with fluid like CSF or proteinaceous content.
[4] Calcification is a collection of calcium in tissues.
[5] Oedema is a swelling tissue caused by excess fluid.
[6] Necrosis is a dead cell in living tissues.
[7] Haemorrhage is a damaged blood vessels resulting in a lot of bleeding in a short time.

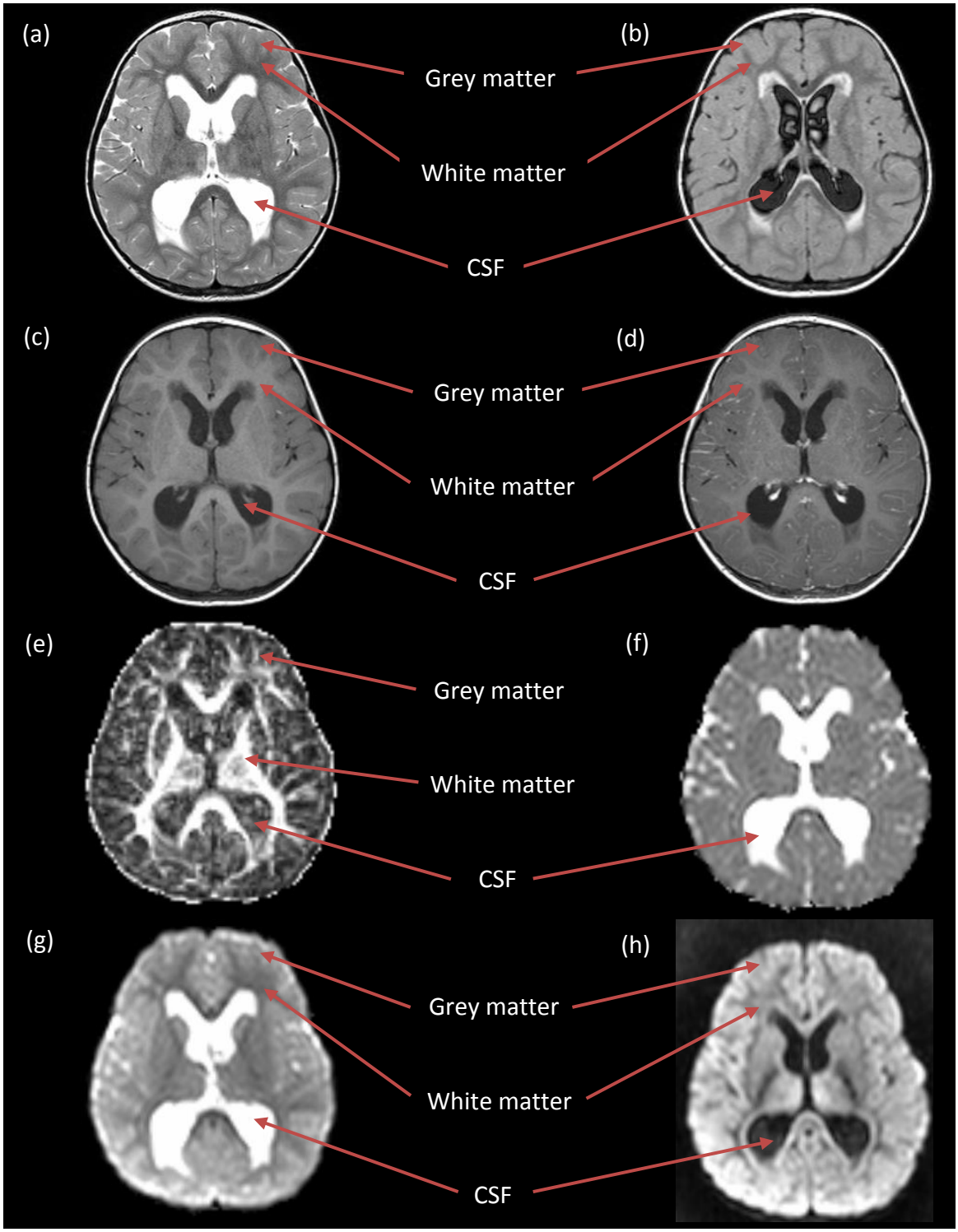


Figure 2.18: Imaging features of grey matter, white matter and CSF obtained from (a) T2, (b) FLAIR, (c) pre-contrast T1, (d) post-contrast T1, (e) FA, (f) ADC, (g) DWI at b = 0 and (h) DWI at b = 1000 s/mm^2 (case MB719).

Grey matter has lower signal intensity and appears darker than white matter on pre-contrast T1 (Figure 2.18 (c)), post-contrast T1 (Figure 2.18 (d)) and FA (Figure 2.18 (e)). Grey matter appears relatively isointense[8] to white matter on an ADC map (Figure 2.18 (f)). CSF is hyperintense[9] relative to normal tissue on T2, ADC and DWI at b=0 images. CSF is hypointense[10] relative to normal tissue on pre-contrast T1, post-contrast T1, FLAIR, DWI at b=1000 s/mm^2 images. CSF appears isointense to normal brain tissue on a FA map.

2.3.2 Medulloblastoma

Medulloblastoma (MB) is a highly malignant tumour (grade IV) of the posterior fossa and prevalent during the first ten years of childhood (74). Medulloblastoma is currently divided into five WHO-defined subsets, which include classic medulloblastoma, desmoplastic/nodular medulloblastoma, medulloblastoma with extensive nodularity, anaplastic medulloblastoma, and large-cell anaplastic medulloblastoma (1). Molecular genetic studies have shown that the five distinct subgroups of medulloblastomas have their own set of characteristics (75). Medulloblastomas are assumed to develop from neural stem cell precursors in the granular cell layer of the cerebellum (76). However, for some subtypes of medulloblastoma, the tumour cell of origin is believed to grow from the superior medullary velum, located at the roof of the fourth ventricle and infiltrate laterally later (74). Although, the exact cellular origin is still controversial, the cerebellum

[8] Isointense is the property of having a very similar signal-intensity between two image regions being compared.
[9] Hyperintense is the property of having a brighter signal-intensity in the region of interest compared to a reference region.
[10] Hypointense is the property of having a darker signal-intensity in the region of interest compared to a reference region.

is still the most common location (94.4%) with most medulloblastomas (75%) arising in the midline, mainly in the inferior and lower vermis (74).

Histologically, the typical features of medulloblastomas are consistent with densely packed, small, round to oval cells with restricted cytoplasm and high mitotic and apoptotic rates (77, 78). The MRI features of medulloblastomas are predominantly homogenous with little necrotic, haemorrhagic, calcification or cystic components (79). Medulloblastomas, containing bands of connective tissue scattered among the small malignant cells, have been classified as desmoplastic medulloblastomas (80). In our study, the majority cohorts include classic medulloblastoma and a few cases of desmoplastic/nodular medulloblastoma. Compared with classic medulloblastoma, the desmoplastic type is less common, tends to occur in older children and adolescents, and is more often lateral in position (80).

Although the MRI features of classic medulloblastoma and desmoplastic/nodular medulloblastoma are visually distinctive, the signal characteristics between the two tumour types was not found to be different on T1-weighted and T2-weighted images (80). The classic and desmoplastic/nodular medulloblastomas are included into the same category for classification in this thesis. Examples of MRI of a classic medulloblastoma and a desmoplastic/nodular medulloblastoma used in this thesis are shown in Figure 2.19 and Figure 2.20.

On T2-weighted images, medulloblastomas are often heterogeneous and have intermediate to moderately high signal. Medulloblastomas appear generally hyperintense to hypointense relative to grey matter and hyperintense relative to white matter, for

example in Figure 2.19 (a) and Figure 2.20 (a). Signal heterogeneity on T2 images is observed in 91% of the lesions and results from intratumoural cystic zones, small blood vessels and/or calcifications (80, 81).

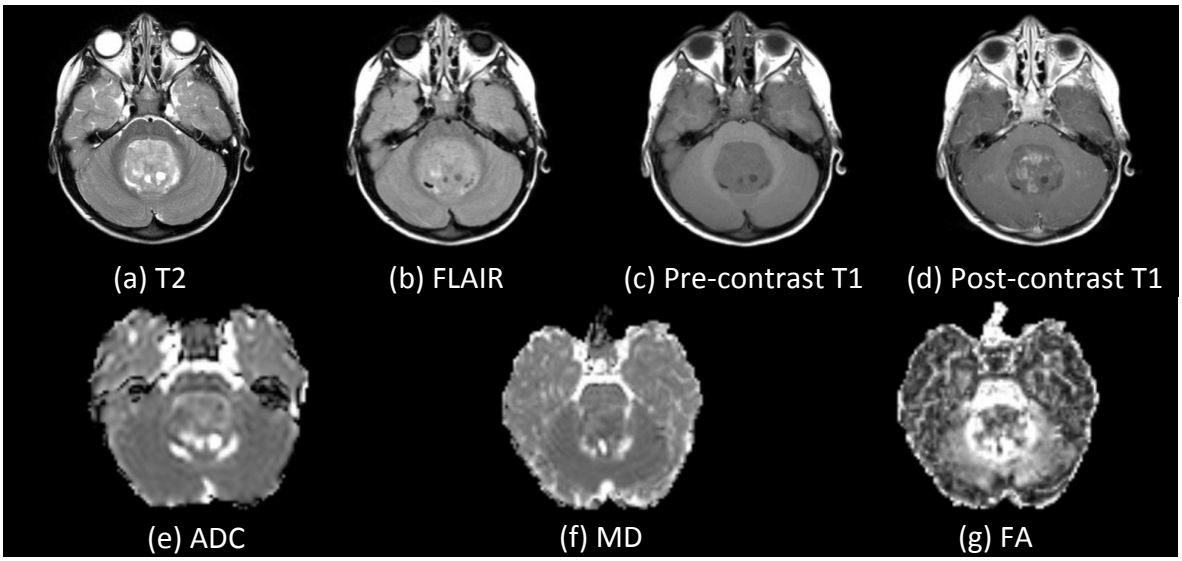


Figure 2.19: Imaging characteristics of a medulloblastoma based on T2, FLAIR, pre-contrast T1, post-contrast T1, ADC, MD and FA (case MB719).

On T1-weighted images, medulloblastomas typically have low to intermediate signal intensity and appear homogenous (Figure 2.19 (c) and Figure 2.20 (c)). They appear iso- to hypointense relative to white matter and predominantly hypointense relative to grey matter (76, 80). Post-contrast T1-weighted images of medulloblastomas (Figure 2.19 (d) and Figure 2.20 (d)), are usually heterogeneous with unlikely regular pattern (82), however, atypical features of homogenous pattern can be found (76). Tumour boundaries on contrast-enhanced images are often less well defined than on the images acquired without contrast. However, post-contrast T1-weighted images present some abnormal tissues clearer than the unenhanced scans which are less likely to show tumour infiltration and only peritumoural oedema (80). On FLAIR images (Figure 2.19 (b), Figure 2.20 (b)), medulloblastomas are generally isointense to normal tissues (83). However, in

some cases, medulloblastomas appear slightly hyperintense to normal tissues as shown in Figure 2.21.

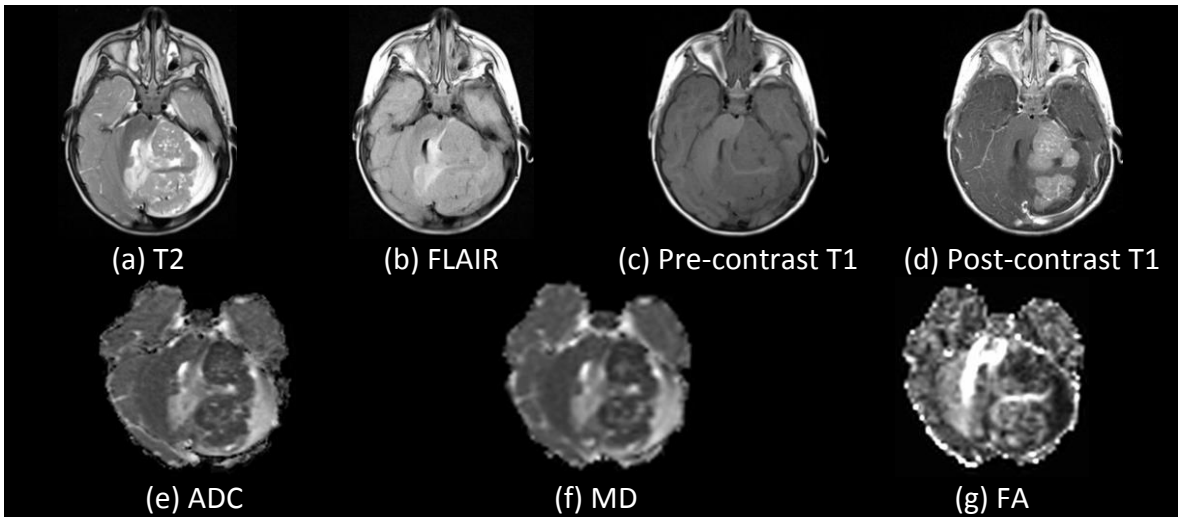


Figure 2.20: Imaging characteristics of a desmoplastic medulloblastoma based on T2, FLAIR, pre-contrast T1, post-contrast T1, ADC, MD and FA (case MB862).

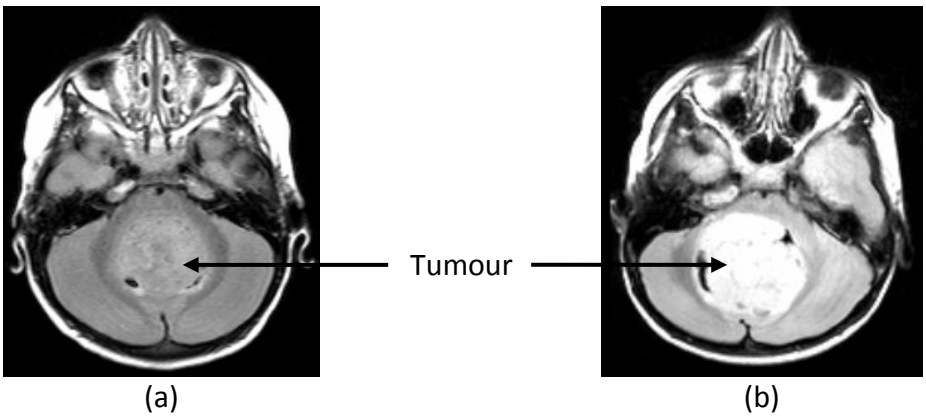


Figure 2.21: Comparison of two FLAIR images of medulloblastoma cases which have different signal intensity level compared to normal tissues. (a) The tumour is isointense to normal tissues (case MB719). (b) The tumour is hyperintense to normal tissues (case MB532).

The diffusion restriction is usually found in high grade lesions. This is due to the high intrinsic cellularity of a tumour, decreased extracellular space and high nucleus/cytoplasm ratio. On DWI, medulloblastomas appear hyperintense to normal brain (84) and exhibit

low signal on ADC maps (79) (Figure 2.19 (e) and Figure 2.20 (e)). Medulloblastoma on MD maps (Figure 2.19 (f) and Figure 2.20 (f)), are similar to that on ADC maps, whereas medulloblastoma on FA maps (Figure 2.19 (g) and Figure 2.20 (g)), show heterogeneous with isointense to hyperintense to normal tissues.

2.3.3 Pilocytic Astrocytoma

Pilocytic astrocytoma is a low grade (WHO I) histologic subtype of astrocytoma and almost never relapses after complete surgical elimination (85). Juvenile pilocytic astrocytoma (JPA) is the most common astrocytic neoplasm in the cerebellum of children. Fewer numbers of pilocytic astrocytomas are found in supratentorial regions, lateral ventricle and brainstem (86). Plenty of pilocytic astrocytoma cells are elongated and bipolar, appearing as 'hairlike' pattern. They are typically well-delineated from the surrounding brain tissue (25). Grossly, the tumours are rarely composed of complete solid tumours (85, 87) but are cystic, with mural nodules and more solid with various cystic central cavities. The cystic compartment is less likely filled with proteinaceous fluid and appears similarly as the signal of CSF.

On a T2-weighted image in Figure 2.22 (a), the signal intensity of pilocytic astrocytoma varies widely from being equivalent to normal CSF intensity to about the same intensity as healthy grey matter, and is generally higher than that observed in medulloblastoma (82, 85) because of the lower cellularity of the pilocytic astrocytoma (9). The tumour is heterogeneous, comprising hyperintense solid mass and hyperintense cystic components relative to normal brain. On a FLAIR image in Figure 2.22 (b), the pilocytic astrocytoma is

generally hyperintense to normal brain tissue (87) and fluid can be differentiated from lesions vividly (88).

On a pre-contrast T1-weighted image in Figure 2.22 (c), the pilocytic astrocytoma appears a wide range of intensity, from about that of white matter to that of CSF. On the pre-contrast T1-weighted image, the tumour is typically composed of heterogeneous region of hypointense solid mass and hypointense cystic component. On a post-contrast T1-weighted image in Figure 2.22 (d), the pilocytic astrocytoma is revealed with heterogeneous lesion with non-enhancing cystic components (87). The enhancement tumour has no definitive pattern. Some atypical pilocytic astrocytomas of the third ventricle shows no significant contrast enhancement (89).

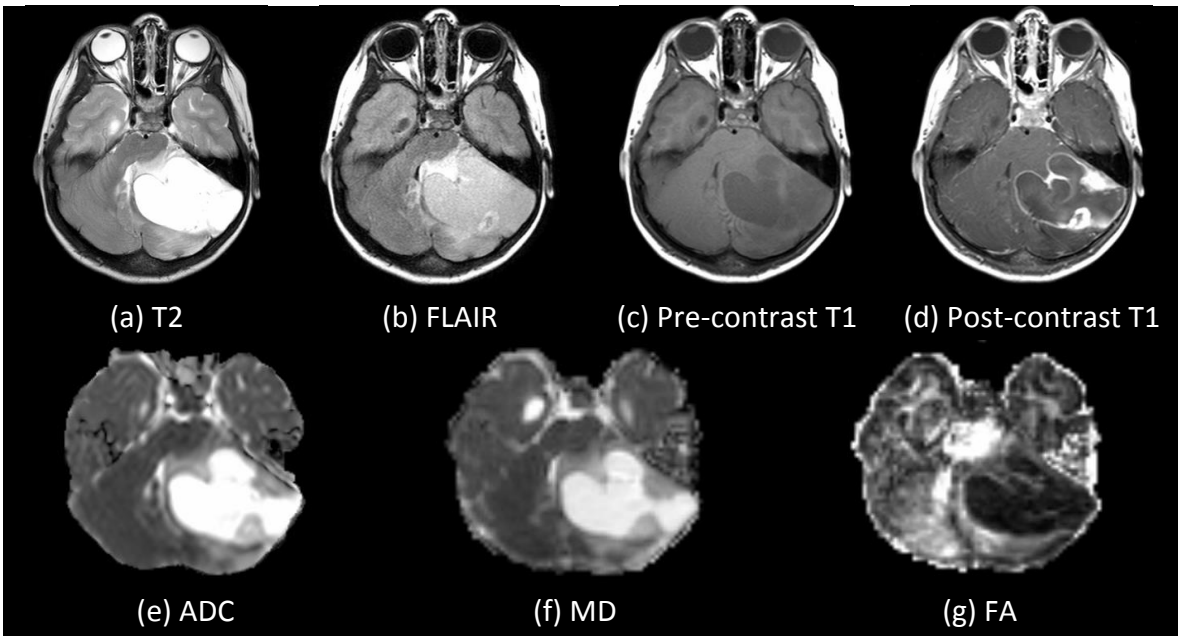


Figure 2.22: Imaging characteristics of a typical pilocytic astrocytoma based on T2, FLAIR, pre-contrast T1, post-contrast T1, ADC, MD and FA (case PA856).

Pilocytic astrocytomas generally have low cell density with relatively large volume of extracellular space. So there is the free movement of water molecules within these

tumours (88) that cause the tumours to appear hypointense on DWI, hyperintense on an ADC map in Figure 2.22 (e) and hyperintense on a MD map in Figure 2.22 (f). On a FA map in Figure 2.22 (g), the solid mass is isointense but cystic component is hypointense relative to normal tissues.

A pilocytic astrocytoma with a dominant solid portion (Figure 2.23) has imaging features that differ from the typical pilocytic astrocytoma, as aforementioned. Its solid mass appears isointense to hypointense relative to normal tissue on T2-weighted images (Figure 2.23 (a), (d)), as well as on ADC maps (Figure 2.23 (c), (f)). The solid mass is hyperintense relative to normal tissues on post-contrast T1 images (Figure 2.23 (b), (e)).

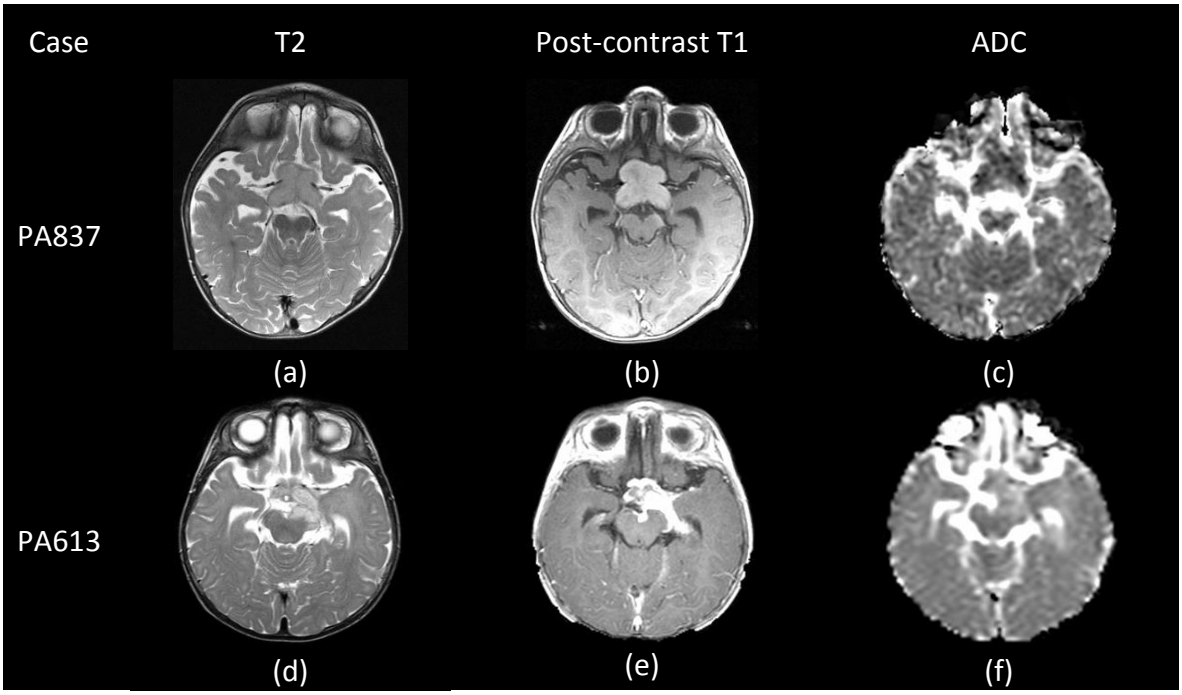


Figure 2.23: Pilocytic astrocytomas with almost exclusively solid and no prominent cystic component.

2.3.4 Ependymoma

Ependymomas are glial tumours mutated from ependymal cells (90) with varying degree of aggressiveness ranging from grade I to III. The WHO classifies ependymal tumours into four types: ependymoma (grade II), anaplastic ependymoma (grade III), subependymoma (grade I) and myxopapillary ependymoma (grade I). Grade II ependymoma and grade III anaplastic ependymoma are difficult to be distinguished by MRI (91). However, they are possibly differentiated if an anaplastic ependymoma has higher incidence of CSF dissemination during diagnosis and poorer prognosis, compared to common grade II ependymoma (10).

Ependymal tumours included in our study were grade II ependymoma and grade III anaplastic ependymoma, because their MRI features are relatively similar and they are most commonly found, compared to other subtypes of ependymomas. Tumours of grade I, myxopapillary ependymoma and subependymoma, were excluded because such cases are limited and they have distinct morphological, biological and demographic features. Myxopapillary ependymoma arises almost entirely in cauda equina region. Subependymoma has different MRI features than ependymoma grade II and occurs more often in an older age group (25).

Ependymomas may arise from three different compartments: infratentorial, supratentorial or spinal region (91). However, ependymomas are mainly located in the posterior fossa and usually arise from the floor of the fourth ventricle in children. The minority are supratentorial ependymomas which are often seen in older children (92). Most supratentorial ependymomas (70%) originate from extraventricular regions in the

cerebral hemispheres, which is believed to have distinct cell of origin from the rests of ependymal cells (10).

Although grade II ependymomas are mostly solid tumours and arise from distinctive ependymal cells in the fourth ventricle and infiltrate through ventricular foramina, they can also be found at the same location with pilocytic astrocytomas in paediatric patients. Both histological tumour types can have similar morphological appearance; for example a pilocytic astrocytoma with no prominent cystic component can appear similar to some subtypes of ependymomas (25). In addition, some paediatric infratentorial ependymomas also have similar or overlapping imaging features with medulloblastomas (91).

Ependymomas are typically heterogeneous masses with areas of necrosis, calcification, cysts with a mural nodule and haemorrhage (93). Water diffusion is restricted within soft tissues for some ependymomas. This is believed to be caused by high cellularity of some lesions. Such characteristics result in a heterogeneous feature on all MRI modalities. Yuh et al. (10) reported that both infratentorial and supratentorial ependymomas generally have hyperintense T2 (Figure 2.24 (a)), and isointense to hyperintense FLAIR (Figure 2.24 (b)), hypointense T1 (Figure 2.24 (c)), hyperintense DWI and hypointense ADC (Figure 2.24 (e)), relative to both grey and white matter (10).

However, Forbes et al. studied the relative intensity of three brain tumour types: pilocytic astrocytoma, medulloblastoma and ependymoma, and represented the relative intensity of ependymoma differently (84). Ependymomas appear hypointense on T2, hypointense on DWI and hyperintense on ADC. From visual inspection, signal intensity of solid portion of ependymoma can range from hypointense to isointense on MD map (Figure 2.24 (f)),

and on FA map (Figure 2.24 (g)), relative to normal brain tissue. On post-contrast T1-weighted image (Figure 2.24 (d)), a mixture of enhanced soft tissues of tumours and non-enhanced components are often present.

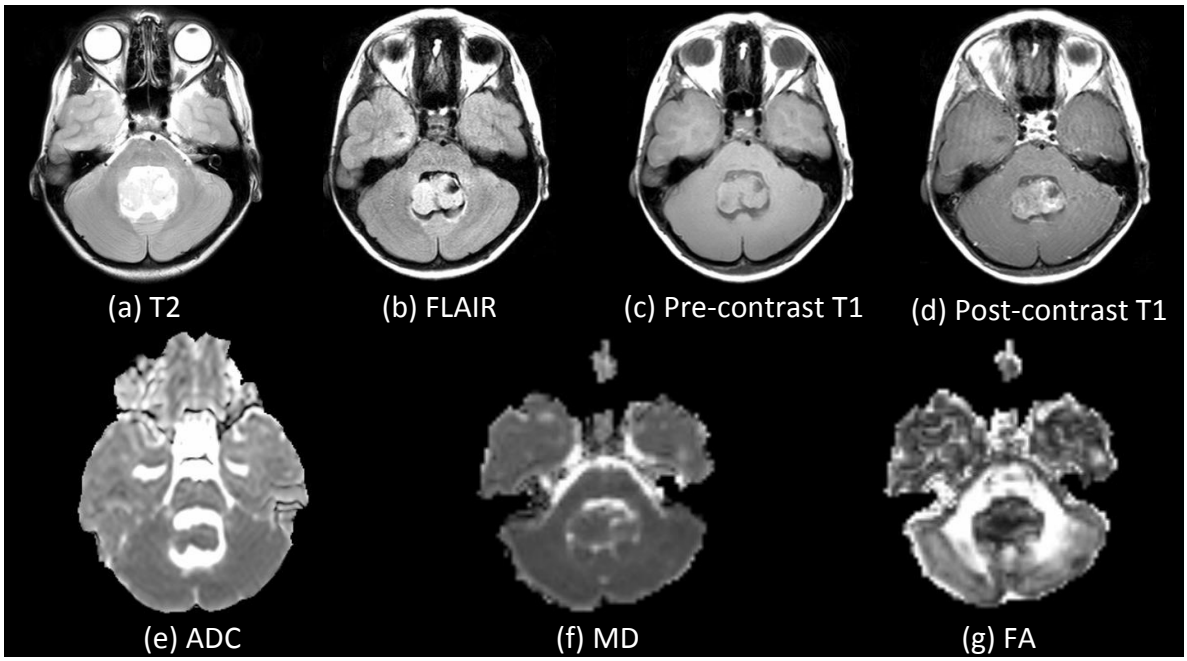


Figure 2.24: Imaging characteristics of a typical grade II ependymoma based on T2, FLAIR, pre-contrast T1, post-contrast T1, ADC, MD and FA (case EP639).

Although infratentorial and supratentorial ependymoma have similar intensity range, supratentorial ependymoma can have a greater tendency of extensive cyst formation and exhibit higher degree of heterogeneity than infratentorial ependymoma (Figure 2.25). Cystic components could have similar signal intensity as fluid on both T2 and FLAIR images. If fluid contains proteinaceous content, signal is hyperintense on FLAIR images. Calcification and old haemorrhage can be found with low signal intensity on all MRI modality. It is observable that diffusivity of both typical ependymomas has a propensity of intermediate intensity between that of pilocytic astrocytomas and medulloblastomas

(10). However, diffusivity values are not an absolute indicator for making histologic diagnoses.

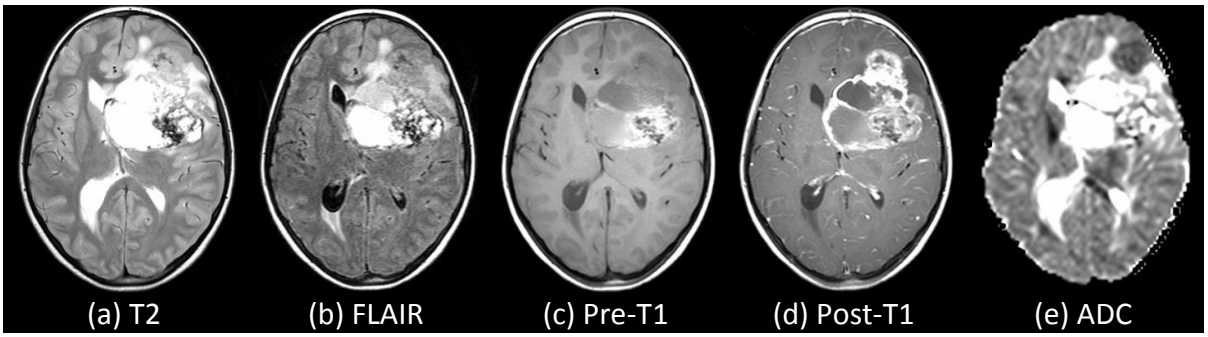
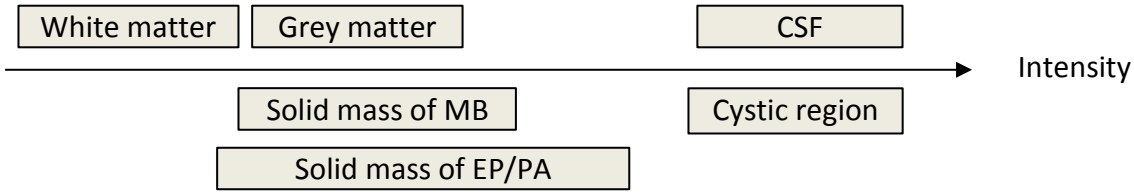
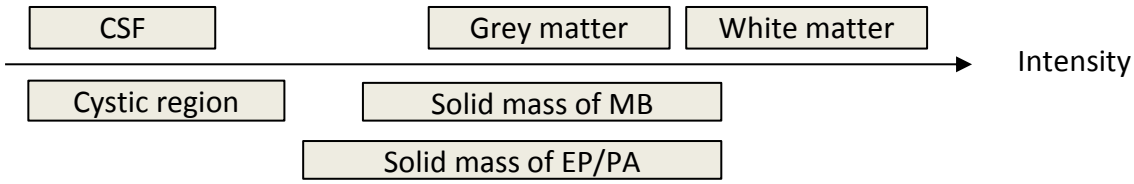


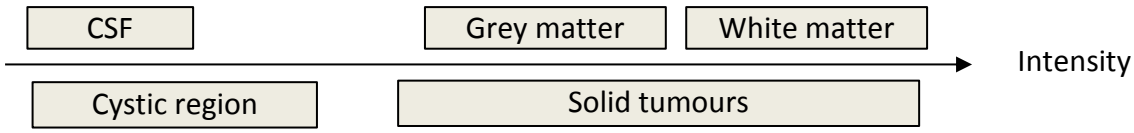
Figure 2.25: Imaging characteristics of a supratentorially located anaplastic ependymoma based on T2, FLAIR, pre-contrast T1, post-contrast T1 and ADC (case EP535).

2.3.5 Discussion

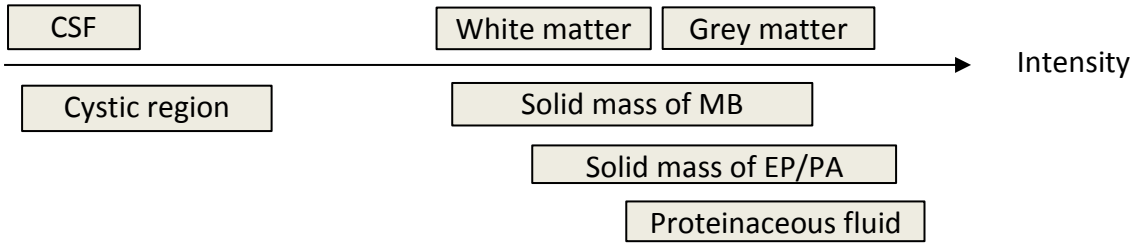
In order to compare the multimodal MRI characteristics of different brain tumour types, the imaging features of the three tumour types acquired from conventional and diffusion MRI sequences are described in Table 2.2. The table also illustrates the typical variation of grey-level intensities of the three tumour types in relation to grey-level intensities of normal tissues: white matter, grey matter and CSF. The estimated range of grey-level intensities of the three tumour types obtained from conventional and diffusion MRI widely overlap. In the literature, the grey-level intensities of the three tumour types PA, EP and MB decrease in that order for ADC images. For other image types, such a pattern has not been reported. However on T2 images MB tends to have lower mean grey-level intensity than PA and EP. This is because MB consists of mostly solid mass and therefore has higher homogeneity. Furthermore, the grey-level intensities of PA and EP based on conventional MRI are noted to have a wider range than MB.

Table 2.2: MRI features for medulloblastoma, pilocytic astrocytoma and ependymoma. Estimated grey-level intensity range is drawn to compare the three tumour types in relation to signal intensity of grey matter, white matter and CSF of each image type. The boxes represent estimated grey-level intensity range.

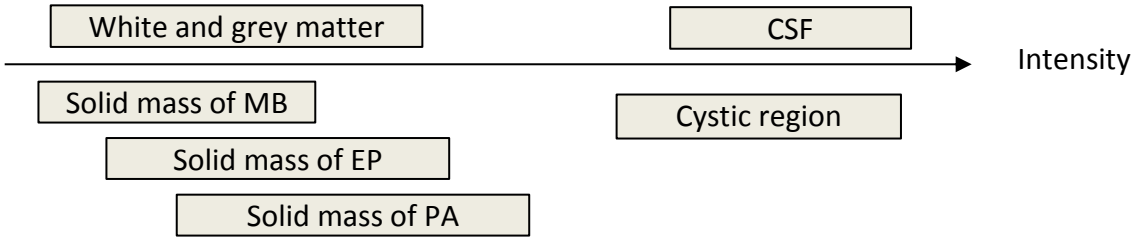
Image type	Medulloblastoma	Pilocytic Astrocytoma	Ependymoma
T2	- MB is heterogeneous with mostly hyperintense solid mass relative to grey matter and white matter.	- PA is heterogeneous with hyperintense solid mass. - Hyperintense cystic components relative to normal brain, and isointense relative to CSF.	- EP is heterogeneous with hyperintense solid mass relative to both grey and white matter. - Cystic components could have similar signal intensity as fluid.
			
Pre-contrast T1	- MB is homogeneous with isointense to hypointense solid mass relative to white matter and mostly hypointense relative to grey matter.	- PA is heterogeneous with hypointense solid mass and hypointense cystic component relative to grey and white matter.	- EP is hypointense relative to both grey and white matter.
			
Post-contrast T1	- MB is heterogeneous or patchy enhancement.	- PA is heterogeneous lesion with non-enhancing cystic components.	- EP is heterogeneous enhanced soft tissues of tumours and non-enhanced components.



FLAIR	- MB is heterogeneous with isointense solid mass relative to normal brain tissue.	- PA is heterogeneous with hyperintense solid mass - Liquid can be differentiated from lesions vividly.	- EP is isointense to hyperintense relative to both grey and white matter - Fluid can be hyperintense if contain proteinaceous fluid.



ADC	- MB is heterogeneous and has lower intensity than PA.	- PA is heterogeneous and higher intensity than MB.	- EP is heterogeneous and intermediate intensity between MB and PA.



Pilocytic astrocytomas, medulloblastomas and ependymomas have different levels of aggressiveness and distinct molecular genetic types, which result in diverse imaging features, e.g. homogeneous and heterogeneous. However, medulloblastomas have a higher propensity to be more homogeneous and reduced free water compared to pilocytic astrocytomas and ependymomas (77, 78, 87). Some cases such as pilocytic astrocytomas without prominent cystic component or ependymomas with large cystic

component can be seen and can be visually misdiagnosed. Pilocytic astrocytomas and ependymomas can be more likely found in supratentorial region than medulloblastomas. Ependymomas can also be found in spinal cord. However, generally, all of these tumour types arise mostly in the posterior fossa rather than the supratentorial region and spinal cord (91, 92).

T2-weighted images are sensitive to tissue pathology and show well-defined tumour delineation (94). The cystic components on T2-weighted images have higher intensity than solid mass. Medulloblastomas on T2-weighted images generally have less cystic components and lower signal intensity compared to pilocytic astrocytomas and ependymomas. T2-weighted images and pre-contrast T1-weighted images show complementary information but pre-contrast T1-weighted images exhibit faint tumour boundary. Post-contrast T1-weighted images can enhance some pathological tissues and increase image contrast on the enhanced region. The cystic and solid portions are hypointense and hyperintense to normal tissues on both pre- and post-contrast T1-weighted images. On post-contrast T1-weighted images, medulloblastomas can have higher intensity than pilocytic astrocytomas and ependymomas. The review of the MRI features of brain tumours on FLAIR images are limited compared to on T1- and T2-weighted images. However, they are claimed to be useful for distinguishing fluid contents. Two main types of fluid show different signal intensity on FLAIR. Fluid like CSF is supressed, while fluid filled with proteinaceous content is hyperintense to normal tissues.

The restricted diffusion, due to the high intrinsic cellularity of tumours, is usually found in high grade lesions. This restricted diffusion results in lower signal intensity on ADC and

MD maps of medulloblastomas, compared to pilocytic astrocytomas and ependymomas. The review of brain tumours on FA maps is limited. However, it has been observed that FA maps exhibit hypointense cystic components compared to heterogeneous solid mass of brain tumour.

The degree of heterogeneity and intensity of tumours, relative to those of normal tissues, cannot completely differentiate the three tumour types through visual inspection. For example, the medulloblastoma, pilocytic astrocytoma and ependymoma in Figure 2.26 have similar morphological shape, with a fluid area, and are found at the same anatomical location. The medulloblastoma images obtain from conventional and diffusion MRI in Figure 2.26 have a large portion of proteinaceous fluid resulting in higher signal intensity on most image types. The pilocytic astrocytoma and ependymoma images obtained from conventional and diffusion MRI in Figure 2.26 have similar morphological shape and may not be visually differentiated. In general, the separation among medulloblastomas with more cystic component; pilocytic astrocytomas with inclusive solid mass; and ependymomas with excessive cystic components are still questionable.

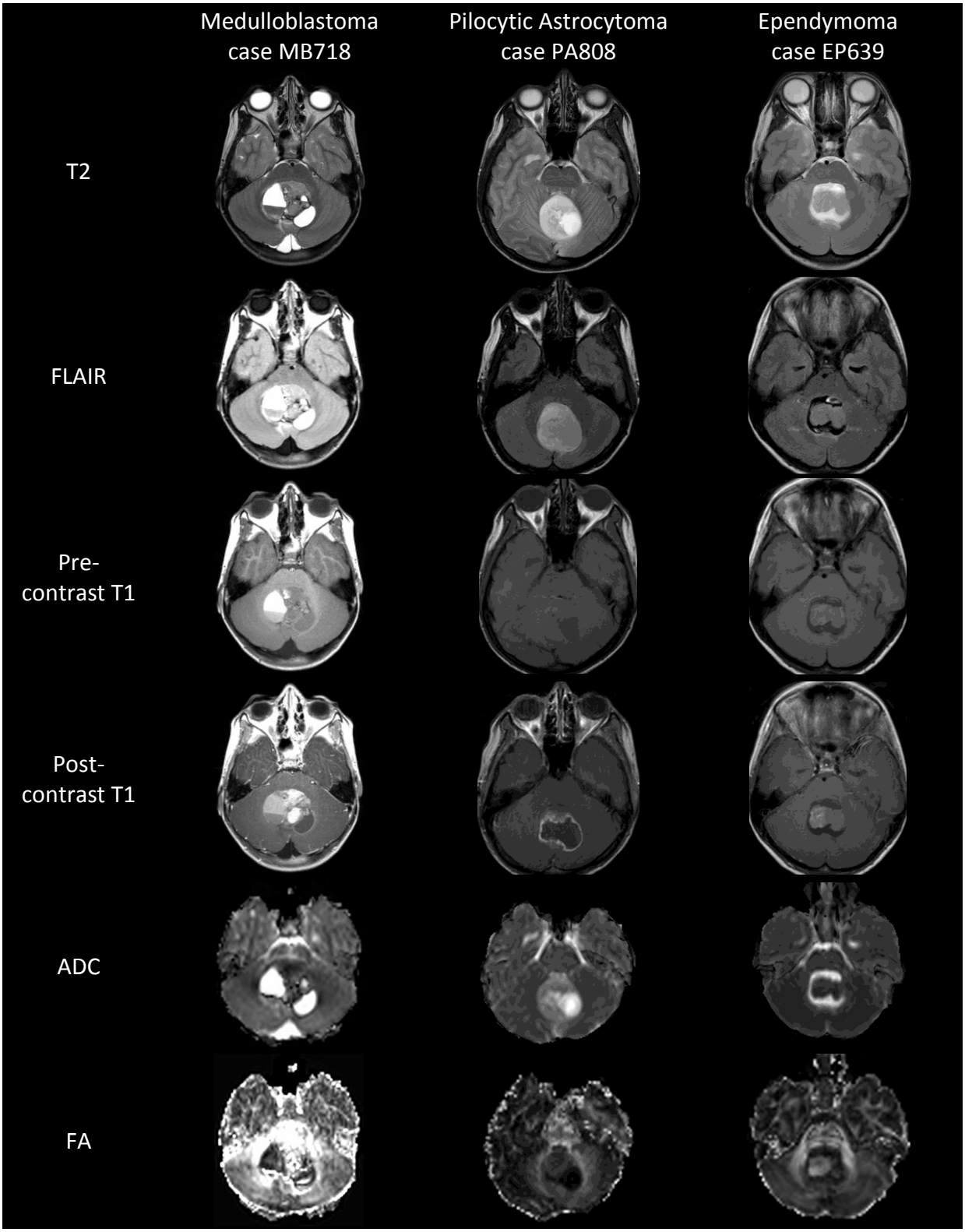


Figure 2.26: Difficulty in differentiating the three tumour types. Imaging features of medulloblastoma, pilocytic astrocytoma and ependymoma are shown based on conventional and diffusion MRI.

2.4 Summary

This chapter reviewed the concepts of MR-based image acquisition and MRI characteristics of the three most common paediatric brain tumour types. The MR-based image acquisition for MR sequences includes T2, T1, FLAIR, DWI and DTI, as well as the reconstruction of diffusion MRI producing ADC, MD and FA maps. Various imaging parameters and their effects on image contrast were discussed. This review provides understanding for MRI features of healthy and pathological tissues, particularly brain tumours, which are the material used throughout this thesis.

The typical MRI characteristics of the three most common paediatric brain tumours: pilocytic astrocytoma, medulloblastoma and ependymoma, as observed on conventional and diffusion MRI were reviewed. In this thesis, subtypes of embryonal, astrocytic and ependymal tumours are examined. Embryonal tumours include classic medulloblastoma (grade IV) and desmoplastic/nodular medulloblastoma (grade IV). Astrocytic tumours include pilocytic astrocytoma (grade I). Ependymal tumours include ependymoma (grade II) and anaplastic ependymoma (grade III). Majority of cases are frequently found in the posterior fossa region of the brain and fewer cases are found in supratentorial region.

The distinctive genetic types of brain tumours contribute to different imaging characteristics, which can appear homogeneous and heterogeneous. The imaging features also depend on imaging sequences, which can enhance abnormal tissues, such as solid mass, cyst, necrosis, calcification and haemorrhage, with different contrast.

The visual diagnosis by using the relative intensity to normal brain tissue and level of heterogeneity is relatively difficult to differentiate these tumour types, with total accuracy. As part of this thesis, texture analysis methods are investigated to better understand such image-based tumour characteristics and potentially provide appropriate imaging markers for computer-aided diagnosis.

CHAPTER 3 THEORETICAL BACKGROUND OF PROCESSING FRAMEWORK

This chapter presents the literature review and theoretical background of the processing framework employed in this thesis, supporting the three main objectives of section 1.2. Multiple methods of neuroimage processing, pattern recognition and statistical analysis are discussed in this chapter.

3.1 Introduction

Computer-aided diagnosis systems have been studied to improve performance of differential diagnosis of brain tumours. Texture analysis based classification of medical images has been shown to produce promising results in several studies, for example in studies of Vlachou et al. (19) and Rodriguez Gutierrez et al. (30). In order to investigate this approach further in the context of childhood brain tumours, a processing pipeline is designed and evaluated. The high-level process flow of the framework has been used in other studies, for example in study of Kassner and Thornhill, where the process involves selection of an ROI, feature generation, feature selection, classification and evaluation (18). This study is focused on TA derived from a single imaging modality. In order to investigate the merits of multimodal TA, we designed a processing pipeline graphically depicted in Figure 3.1. The process pipeline contains data pre-processing step which enables multimodal TA-based classification. In order to select the most suitable techniques for each step in the pipeline, a literature review is conducted. In the following

sections we present the reasoning behind specific choices and a concise theoretical background of the chosen methods.

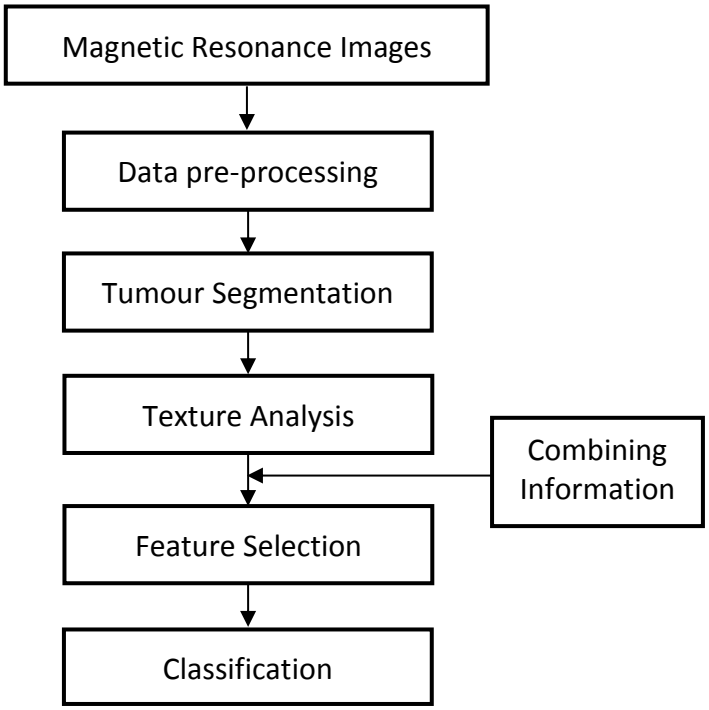


Figure 3.1: A general framework of multimodal MRI based texture analysis for classification of childhood brain tumours.

3.2 Pre-Processing Analysis

Four methods involve artefact removal, skull stripping, image registration and intensity normalisation are reviewed and considered as pre-processing analysis, described in the following sections.

3.2.1 Artefact Removal

Artefacts are undesirable features which obscure or otherwise hinder the extraction of useful image data. Artefact removal is important to improve the SNR and to produce

sufficiently clean input data for further analysis. Generally, artefacts related MRI can be generated by various sources and broadly divided into three categories: motion, field inhomogeneity and digital imaging (41). Motion artefacts can be produced from disordered muscular action, so-called involuntary movement, patient motion and physiological movement such as blood flow. Motion artefacts commonly appear as ghosts or blurring along the phase-encode direction, however, they can be potentially corrected by B-matrix reorientation for DTI data (68) and image registration methods (95). Field inhomogeneity produces off-resonance, any signal which has a frequency different from the Larmor frequency of a nuclear spin in a perfectly uniform static magnetic field. The off-resonance can be caused by hardware imperfection due to, for instance, a non-optimised setting (96) and susceptibility effects within the human body, for example air in the sinuses. This artefact typically causes geometric and intensity distortion on an image (Derek and et al., 2001). The third artefact type is induced by digital imaging artefacts, generating for example phase wrap-around artefacts. The phase wrap-around artefacts are present when the anatomy is larger than the field of view (FOV) and is recorded at the opposite edge of the scan (41).

In this thesis, the emphasis of artefact removal is on the motion and field inhomogeneity artefacts. Correction of digital imaging artefact due to an inappropriate used FOV is beyond our scope of study because we can only access to the post-imaging data. The motion artefacts are corrected by using co-registration of input images (including MRI and diffusion MRI) onto T2-weighted images. The field inhomogeneity is corrected for cases of diffusion MRI dataset which are commonly acquired with EPI (97, 98) and is susceptible to

the off-resonance artefacts influenced by such as B_0 field inhomogeneity and the eddy current effects (99).

B_0 inhomogeneity affects EPI by producing geometric distortion which is seen as pixel shifts in the read-out direction and particularly in the phase-encode direction (96). The B_0 inhomogeneity effects can be removed by using a B_0 field mapping approach and image registration. Comparing the two correction methods, Metaxas et al. (98) showed that B_0 field mapping was superior than a deformable registration method. B_0 field mapping requires a fieldmap (phasemap) to unwrap phase discontinuities and modulate the k-space data in Fourier domain (100). The field mapping method can only correct B_0 field inhomogeneity, but not the off-resonance induced by eddy currents.

Eddy currents are induced currents, which are generated when large gradient pulses are switched on and off rapidly, such as those used for diffusion MRI (101). Eddy currents induce geometric distortions which appear as image shearing in the read-out gradient field direction, scaling in phase-encode gradient field direction and bulk shifting (102). The significance of eddy current-induced geometric distortion depends on the direction and strength of the applied diffusion field gradients, for example different b-values applied in different gradient direction. The distortion can be partially corrected by using geometric registration to a reference volume (Jezzard et al., 1998). Other techniques have been implemented for eddy current-induced artefacts, for example using an iterative cross-correlation algorithm (103), using a twice-refocused spin echo (104), using bipolar diffusion-weighting schemes (105), and using interleaved echo planar techniques instead of single shot EPI (102).

3.2.2 Skull Stripping

Whole brain extraction, skull-stripping, or brain/non brain segmentation is a process of separating brain tissue (e.g. grey matter, white matter, CSF) from non-brain (e.g. skull, scalp and dura). Accurate whole brain extraction is an important initial step in image processing for inhomogeneity correction, image registration, intensity normalisation, and tissue classification (106-108). Various automatic segmentation methods have been implemented based on three main approaches that are pixel classification, boundary-based segmentation and surface-based model (109). However, hybrid segmentation approaches are more commonly implemented because they can achieve a better result than a single algorithm approach. A number of software tools include a skull stripping algorithm as part of their packages, reviewed as follows.

Shattuck et al. developed Brain Surface Extractor (BSE) as part of the BrainSuite (110). BSE incorporates two main algorithms: edge detection and morphology. After anisotropic diffusion filtering, a 2D Marr-Hildreth edge detector (convolving an image with the Laplacian of the Gaussian function) is initially applied to obtain a brain edge map. The brain surface is consequently obtained from the edge map by using mathematical morphology.

Smith implemented Brain Extraction Tool (BET) as part of the FSL package (111). BET integrates an intensity-based estimation of the brain/non-brain threshold and a deformable model to segment brain from non-brain elements. The method approximates the minimum and maximum intensity values of brain volumes and subsequently applies a

deformable model to shape the brain surface by using a smoothness criteria and a local intensity threshold.

Segonne et al. implemented Hybrid Watershed Algorithm (HWA) as part of the FreeSurfer tool (112). HWA combines watershed algorithms and deformable surface models. The watershed algorithm provides an initial estimate of the brain volume and incorporates with the deformable surface model to refine the brain surface.

Boesen et al. developed a method called Minneapolis Consensus Strip (McStrip), which uses a hybrid algorithm incorporating three procedures: warping to a template, intensity thresholding and edge detection (113). The study reported that McStrip consistently outperformed SPM2, BET and BSE. However, the McStrip method requires much longer processing time than the other techniques and may not be suitable for a large number of cases.

Apart from the well-known skull stripping methods, several other algorithms have also been developed, for example a hybrid approach of expectation maximisation and geodesic active contours (114), a skull stripping based on graph cuts (107) and a Multi-Atlas Propagation and Segmentation (MAPS) (108). However, these techniques have not been widely adapted in the field of medical image analysis. A well-known skull removal technique, which has been used in a number of studies, is the BET method. The BET method is considered as state of the art and is part of the FSL package. Therefore, it is convenient to employ the BET method in the processing pipeline and it also provides fast processing time compared to other methods.

3.2.3 Image Registration

Image registration is an important process for removing motion artefacts and incorporating multimodal image analysis in this study. The registration aligns an image to a template of the same object acquired, either at different times, from different scanning directions or by different imaging devices (115). Basic components of registration algorithms comprise three parts: 1) transformation, 2) similarity measures and cost function, and 3) optimisation (116, 117), as described in the following sections.

3.2.3.1 Transformations

Transformation specifies a mapping function in order to overlay a floating image (X) on a reference image (Y) geometrically. The transformation model provides two main operations: control the spatial mapping and interpolate missing features between reference and floating images (116). The complexity of transformation in terms of elasticity, reliability and computational load is explained by the degrees of freedom (DOF). DOF is the number of independent transformation in a specified mapping function. For example, translation in 3D space encompasses 3 independent translation, which is equivalent to 3 DOF transformation model (117). The transformation model can be classified into three main types: rigid, affine and non-linear transformation.

Rigid transformations are linear operations, which preserve the shape and size of an object. These techniques are suitable for intramodal imaging data of the same subject that has no distortion and no anatomical changes (118). A rigid transformation comprises 6 DOF: three rotations and three translations, and provides the fastest calculation.

Affine transformations offer higher DOF of linear transformations than rigid transformations. They allow linear coordinate changes up to 12 DOF including rotation, translation, scaling and skews (shear). Affine transformations are typically used as an initialisation to nonrigid transformations, either for primary approximation of location, or compensation for geometric image distortions affected by eddy currents in diffusion MRI, for instance (119).

Nonrigid, non-linear, deformable or elastic transformations are more commonly used for images of different subjects, images with distortions during acquisition or images with actual physical differences such as biological changes (116). The non-linear transformations have at least 12 to millions of DOF. The choice of DOF affects the computational load and the accuracy of the registration. Generally a low DOF or affine transformations are more useful than a very high DOF because they are less sensitive to artefacts and poor image quality, hence a more robust model. A high DOF may be more beneficial for an image, having moderate to high resolution with good contrast and little to no artefacts (117).

An interpolation is considered as part of the transformation process and is used to fill the missing values between the reference and floating images. The interpolation method is calculated by convolving the image data with a continuous kernel. The kernel functions, which have been commonly used, are for example trilinear, spline and sinc kernels. Sinc interpolation was showed to provide more accuracy than the trilinear approach, as recommended by Jenkinson et al. (120). Therefore sinc interpolation is used with a default window width 7. The optimisation adopts a multi-resolution approach with a local

optimisation method (95, 121, 122). Two methods for dealing with the local minima problem are cost function apodization, which is used to reduce or eliminate small-scale dips, and the hybrid global-local optimisation technique. The details of the optimisation are beyond the scope of this thesis.

3.2.3.2 Similarity and Cost Functions

Similarity and cost functions measure the quality of a transformation model, orientating two images on the same spatial location. Cost functions can be broadly divided to geometric- and intensity-based cost function. Intensity-based cost functions offer higher accuracy and reliability than geometric-based cost functions (120). Intensity-based cost functions can be categorised by intramodal and intermodal registration. For registration of intramodal imaging, the most commonly used functions are least squares and normalised correlation. For registration of intermodal imaging, the most commonly used functions are woods, mutual information, normalised mutual information, and correlation ratio.

In FLIRT, the correlation ratio (123) is a default similarity function, while normalised mutual information is an alternative cost function. The correlation ratio (CR) and normalised mutual information (NMI) are mathematically defined in equation (3.1) and (3.2) respectively.

$$Correlation\ Ratio = \frac{1}{Var(I_B)} \sum_i \frac{n_i}{N} Var(I_{Bi}) \tag{3.1}$$

$$Normalised\ Mutual\ Information = \frac{H(I_A, I_B)}{H(I_A)\ +\ H(I_B)} \tag{3.2}$$

where I_A and I_B represent a set of image intensity of image A and B; I_{Bi} is the ith set of intensities of image I_B at positions where the intensity in I_A is the ith intensity bin; $Var(I_{Bi})$ is the variance of the I_B in area i; $Var(I_B)$ is the variance of set I_B; $N = \sum_i n_i$; n_i is the number of elements in the set I_{Bi}; $H(I_A,\ I_B) = -\sum_{ij} \frac{n_{ij}}{N} log\left(\frac{nij}{N}\right)$ where n_{ij} is the number of voxels that are assigned to the bin pair (i, j); The marginal entropies $H(I_A)$ and $H(I_B)$ are defined similarly, but using the individual image histograms rather than the joint histogram.

3.2.3.3 Optimisation

After a cost function is chosen, the transformation (T^*) is optimised by minimising the cost function as mathematically defined by equation (3.3) (120).

$$T^* = \arg\min_{\mathrm{T} \in \mathrm{S_T}} \mathrm{C}(I_B, \mathrm{T}(I_A)) \tag{3.3}$$

where S_T is the transformable space, I_B is the reference images, I_A is the floating images, $\mathrm{C}(I_B, \mathrm{T}(I_A))$ is the cost function and $\mathrm{T}(I_A)$ represents the transformed image I_A by the transformation T.

3.2.4 Intensity Normalisation

The integration of multi-centre MRI data to acquire sufficient study cases produces disparities in the dynamic intensity range of brain tissue. In addition, tissue intensity behaviours of brain MRI volumes can also vary significantly due to multiple sources of variation, such as different manufacturers and scanner-models, different pulse sequence parameters, field inhomogeneity, heterogeneity of tissues and stage of disease progression. To deal with intensity variations of raw MR images, two main approaches:

intensity inhomogeneity correction and intensity normalisation, have been mentioned (124). Intensity inhomogeneity is corrected using a mathematical model, as well as controlling the signal acquisition and instrument, such as the MR pulse sequence and the imaging coils (125). In this thesis, we can only access post-imaging data; therefore only intensity normalisation is considered here.

Standardising intensity scale is useful in texture analysis to remove dependence on the individual MRI setting, affecting gain and image contrast (26). It is also useful to generalise the relative behaviour of various tissue types across different MRI volumes in the presence of pathology (124). In classification system, Collewet et al. showed that supervised automated tissue classification methods rely strongly on intensity normalisation (126). They evaluated the effect of MRI acquisition protocols and grey-level normalisation methods on texture-based classification. Their results showed that the classification outcome is affected by the normalisation methods and the acquisition protocol. In the case of no normalisation methods, the level of classification errors is dependent on the MR acquisition protocols. Whereas, applying normalisation improves classification accuracy and the relationship with the acquisition protocols was negligible.

Various intensity normalisation methods have been proposed to address the issue of standardising the tissue intensity ranging across MR data. Shah et al. provided a comprehensive review of intensity normalisation techniques based on MR images of multiple sclerosis (124). Early work on intensity normalisation produced the dynamic histogram warping method (127) applied on a pair of stereo image. Although this technique was reported to produce satisfactory results for stereo images, it is less

intuitive to apply on medical imaging with non-uniform tissue composition and artefacts (128).

Wells et al. proposed an adaptive segmentation of MRI data together with intensity inhomogeneity removal (129). This method requires two main parameters: tissue class probability and gain field estimation. Then, the method supplies the parameters into an iterative Expectation-Maximisation method to refine the parameters for estimating classification and intensity normalisation. The acquisition of class conditional probability density functions and gain field could be difficult and time consuming.

Sled et al. proposed a method for correcting intensity nonuniformity in MRI data with a non-parametric method (130). This method does not require parametric model of tissue intensities or segmentation of contiguous regions.

Nyul and Udupa proposed a method consisting of two stages: a training stage and transformation stage (131). The training stage is initially performed to obtain the standard scale from the training dataset. Subsequently, the transformation stage maps a histograms of original image intensity to the standard scale in a piece-wise linear manner. This two-stage method has some significant advantages over other methods. It is both easier to customise to various anatomical regions and fast in practice.

Hellier proposed a brain intensity normalisation of MR images by using a Gaussian mixture model computed through an Expectation-Maximisation algorithm (132). Mean intensity of five main anatomical tissues (background, CSF, grey matter, white matter and mixture of fat and muscle) was computed and mapped between source and target image.

However, pathological tissues may not fall into the five tissue categories and therefore it may be difficult to standardise the intensity.

Chirstensen utilised even-order histogram derivatives to calculate characteristic values of white matter in various image modalities (133). A 3D model is rendered by interpolating the images. The images were de-noised improving the SNR by using histogram analysis and a box-car filter. The characteristic of white matter intensity was determined from the minimum of the 2nd and 6th derivatives and the maximum of the 4th derivative.

Weisenfeld and Warfteld estimated a multiplicative correction field to match a template histogram to a reference model density (134). This method uses the Kullback–Leibler divergence (135) to optimise global statistics of template histogram and maintain local feature contrast. This approach requires pixelwise correction and yields slow processing time.

Jäger et al. utilised non-rigid registration to map multimodal MRI to perform intensity normalisation (136). They used the probability densities of tissue types as images and applied a distance measure based non-rigid registration to the joint histograms. The method was applied on multimodal images and required the histogram to come from at least two imaging modalities. Although, the method gave promising results, performing non-rigid registration operations over multimodal images makes this method slower than other techniques that are independent to image pre-processing and operate on single modality.

From the aforementioned techniques, Shah et al. (124) concluded that the decile-based piecewise method of Nyul and Udupa (131) has obtained a wide acceptance among these,

mainly as a result of the ease of application and speed of execution without loss of accuracy. In the following sections, the theoretical foundations of the decile-based piecewise linear transform and linear normalisation methods are presented as both methods are utilised in this thesis.

3.2.4.1 Linear Normalisation

Linear normalisation is a simple technique that transforms original intensity to a standard scale. Four quantities: the minimum and maximum values of the standard scale as well as the minimum and maximum intensity of the image being normalised, are considered. The minimum and maximum intensity of images are usually considered as outlier intensities caused by artefacts (131). The intensity at specific low and high percentiles of the overall image intensity range is commonly used as the minimum and maximum robust intensity, $[I_{min}, I_{max}]$. Pixels having intensity less than the robust minimum are set to zero, while pixels having intensity greater than the robust maximum is set to the maximum standard scale. Then, the rest of the pixels, having intensity values between the robust minimum and maximum value, are scaled to I_n as shown in equation (3.4). Diagram of linear normalisation is shown in Figure 3.2.

$$I_n = K_n(I_c - I_{min})/(I_{max} - I_{min}) \tag{3.4}$$

where K_n is the maximum intensity value, the minimum intensity value is zero, I_c is the original intensity being considered.

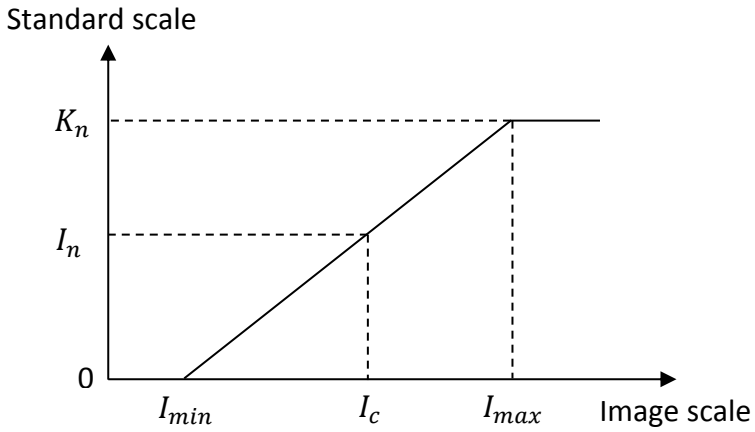


Figure 3.2: Intensity normalisation based on linear normalisation method.

3.2.4.2 Decile-Based Piecewise Linear Transform

This method consists of two stages. One is a training stage to find values for the standard scale. Another is a transformation stage that maps the histograms of candidate volumes to the standard histogram scale in a piece-wise linear manner, shown in Figure 3.3. The two-stage intensity normalisation algorithm is summarised as follows:

In the training stage, the intensity landmark is set to p_L.

$$p_L = [p_{min}, p_{10}, p_{20}, p_{30}, p_{40}, p_{50}, p_{60}, p_{70}, p_{80}, p_{90}, p_{max}] \tag{3.5}$$

p_L is calculated for each training case. The minimum and maximum percentile value of the overall intensity range is obtained at specific percentiles to avoid the tail of histogram. The other intensity landmark is obtained at $j \in \{10, 20, 30,\ldots, 90\}$ percentile. The average of the mentioned intensity is calculated and is defined as standard scale S.

$$S = [s_{min}, s_{10}, s_{20}, s_{30}, s_{40}, s_{50}, s_{60}, s_{70}, s_{80}, s_{90}, s_{max}] \tag{3.6}$$

In the transformation stage, the intensity within a region of interest $I \in [p_{2j}, p_{1j}]$ is mapped to the standard scale derived from the training stage. A transformed intensity I_n is defined mathematically as:

$$I_n = s_1 + \frac{s_2 - s_1}{p_{2j} - p_{1j}}(I - p_{1j}) \tag{3.7}$$

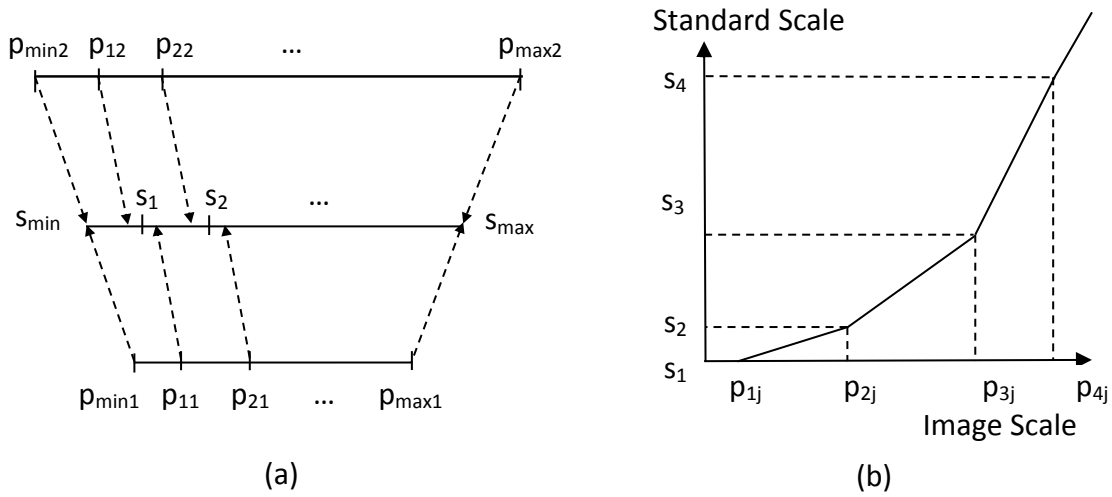


Figure 3.3: Intensity normalisation based on piecewise linear transform method (a) training stage and (b) transformation stage.

3.2.4.3 Comparison of Intensity Normalisation Methods

The two intensity normalisation methods and original data are compared to examine the efficiency of the intensity normalisation methods. The methods are compared by measuring the mean and standard deviation of intensity within the same tumour type of individual MRI types. The mean intensity of entire sample for individual MRI is normalised to zero mean and one standard deviation. Subsequently, the power of discrimination is measured by sum of mean difference square (SMD), shown in equation (3.8).

$$SMD = \sum_{a} (a - \mu)^2 \tag{3.8}$$

where $\mu = (\mu_{MB} + \mu_{PA} + \mu_{EP})/3$, $a \in \{\mu_{MB}, \mu_{PA}, \mu_{EP}\}$, μ_{MB}, μ_{PA} and μ_{EP} are the mean of intensity of medulloblastomas, pilocytic astrocytomas and ependymomas respectively.

3.3 Brain Tumour Segmentation

In image processing, segmentation is the process of partitioning a digital image into homogeneous sub-regions. The goal is to detect the ROI from background or unrelated regions. Segmentation techniques can be divided into manual, semi-automatic, and fully automatic segmentation based on the level of user interaction. Manual segmentation relies on the domain-knowledge of the user, and therefore the resulting output from a recognised expert is commonly used as the ground truth or the gold standard. However, this technique is labour intensive, less likely to be reproducible as well as subject to intra-observer and inter-observer variability (137, 138). Fully automatic segmentations require no operator interaction and can potentially achieve faster processing time. However, they are likely to perform less satisfactorily on medical images because of the complexity and heterogeneity of anatomical texture (139, 140), which would be easier to identify visually but not automatically. To deal with these issues, semi-automatic segmentation methods, integrating operator's supervision with computer-aided analysis to define an optimal region of interest, is investigated.

Classically, segmentation algorithms for medical images can be categorised into four main categories: thresholding, edge-based, region-based and classification approaches (141). Thresholding separates pixels into different classes, depending on their grey-level intensity and a number of threshold values. Edge-based methods detect edges or boundaries of different regions by considering rapid changes in intensity near edges,

whereas region-based methods consider groups of pixels to form a homogenous region and partition an image into sub-regions. In the classification approach, an image can be partitioned according to supervised or unsupervised criteria (142). A supervised approach assigns weighting values to pixels of training images and partitions a test image based on a training model, whereas an unsupervised approach does not require training data but uses features within an image to determine classes. The use of more than one segmentation techniques, referred to as hybrid methods, can improve segmentation performance because there is no ideal method that is well suited for all types of images nor can all methods perform equally well for a specific type of image (143). In the following section, some of the popular segmentation methods are reviewed, with a particular emphasis on brain tumour segmentation.

Grey-level thresholding is considered as one of the simplest image segmentation methods and is also adapted in a number of segmentation methods. Edge detection algorithms offer edge detection operators such as Canny (144) and Sobel (145), which are used to reveal boundaries of objects in different regions. Both thresholding and edge detector algorithms provide results with a fast processing time. However, these methods are inappropriate for an image without apparent peaks of intensity representing distinctive regions and may not provide contiguous segments because there is no consideration of spatial information (143). There is still a need for post-processing, such as edge linking or morphological operations in order to connect lines or remove holes to form a complete boundary. It is also less likely to be suitable for segmenting MR images of brain tumours, which are often heterogeneous in structure (1).

Clustering techniques, categorised as a classification approach such as an unsupervised K-means method (146), are used to classify pixels in an extracted feature space. Nevertheless, the use of variance to measure the cluster scatter and its sensitivity to noise may not be ideal for MR images (147). Watershed segmentation (148) is a region-based technique that customises image morphology and requires at least one marker to partition an image. When the markers are set, the region can be grown by using a morphological watershed transformation (149). This method produces closed and adjacent contours, however, it often presents results with the problem of over-segmentation when an image is highly contaminated or has low SNR (150, 151). Although a region merging technique has been proposed to resolve this issue, it could lead to under-segmentation of homogeneous sub-regions (150).

The deformable model, active contour or snake approach, an advanced edge-based algorithm, is one of the recommended methods used in medical image segmentation (147, 152, 153). The Gradient Vector Flow (GVF) snake method, as proposed by Chenyang and Prince (154), is a preferred technique to the traditional snake method (155) because it is less sensitive to initialisation conditions and is more capable to move into concave boundary regions (156). This method was reported to perform satisfactorily on simple structures with a reasonably well defined boundary around an object. Nonetheless, due to the structural complexity observed on MR images of brain tumours, an active contour can progress towards the wrong boundary (157). It is usually preferable to place seeds (selected coordinates) close to the true boundary of the tumour region, however, this can be time consuming (32, 153).

Graph-based segmentation uses the similarity and dissimilarity of pixels to partition an image into different set of patterns. Given an image, a pixel or a group of pixels form a node and edges define relationships between neighbouring pixels. Each partition of the nodes (pixels) output from this algorithm denotes a segment in the image. The multiscale normalised cuts (MNcut) method, as introduced by Cour et al. is a recently developed graph-based segmentation technique and can be considered as a hybrid method of edge-based and region-based approach (158). This method adopts the normalised cut method (159) as a partitioning framework and offers the advantage of detecting coherent regions with a faint boundary, which is hypothesised to be appropriate for brain tumours since they are often infiltrative, thus appearing with faded boundaries on MRI.

In the following sections, the theoretical foundations of the GVF snake and MNcut methods are presented as both methods are utilised for the segmentation of brain tumour images in this thesis.

3.3.1 Gradient Vector Flow Snake

Snakes are active contour models that progress toward an image boundary by minimising an energy function influenced by internal and external forces. The internal forces are defined by the active contour model and the external forces are derived from the image itself. These forces are designed such that the snake can conform to an object boundary or other desired features within an image. Gradient Vector Flow (GVF) approach defines a parametric active contour model and improves the capture range of the boundary better than the conventional snake technique by using the GVF field as the external force. This GVF snake method is further discussed below.

Given an image having coordinates $(x, y) \in \mathcal{R}^2$, a traditional two-dimensional snake defines a parametric curve function as $\boldsymbol{c}(s) = [x(s), y(s)]$, s ∈ [0,1]. The contour moves through an image under the influence of internal and external forces to minimise the energy function as defined in equation (3.9).

$$E = \int_0^1 \frac{1}{2}[\alpha|\boldsymbol{c}'(s)|^2 + \beta|\boldsymbol{c}''(s)|^2] + E_{ext}(\boldsymbol{c}(s))ds \tag{3.9}$$

where α and β are coefficients governing the snake's tension and stiffness respectively. The tension controls the stretching and stiffness provides rigidity of the elastic curve. The $\boldsymbol{c}'(s)$ and $\boldsymbol{c}''(s)$ represent the first and second derivatives of the parametric curve function $\boldsymbol{c}(s)$ with respect to s, where s is the parametric domain. The external energy function $\boldsymbol{E_{ext}}$ is derived from a grey-scale image, explained as follows.

Given a grey-level image, $I(x, y)$ is a function of position variable (x, y). The external energy function is constructed to provide the deformable contour by blurring the image with the Gaussian function, as shown in equation (3.10).

$$E_{ext}(x, y) = -|\nabla(G_\sigma(x, y) * I(x, y))|^2 \tag{3.10}$$

where $G_\sigma(x, y)$ is a two-dimensional Gaussian function with standard deviation σ, $*$ denotes the convolution, ∇ is the gradient operator of an edge map. A larger σ results in a more blurry image and potentially increases the capture range of the target boundary. A deformable contour that minimises energy function E must satisfy the Euler equation, as shown in equation (3.11).

$$\boldsymbol{F_{int}} + \boldsymbol{F_{ext}} = 0 \tag{3.11}$$

The internal force $\boldsymbol{F_{int}} = \alpha \boldsymbol{c}''(s) - \beta \boldsymbol{c}''''(s)$ maintains the shape of curvature and prevents the curve from stretching and bending. The external potential force $\boldsymbol{F}_{ext} = -\nabla \boldsymbol{E_{ext}}$ forms the deformable contour to progress toward the image contour. The equation (3.11) can be rewritten as in equation (3.12).

$$\alpha \boldsymbol{c}''(s) - \beta \boldsymbol{c}''''(s) - \nabla \boldsymbol{E_{ext}} = 0 \tag{3.12}$$

To find a solution to equation (3.12), the snake is dynamically deformed by altering $\boldsymbol{c}$ as a function of time t and curvature s. Then the partial derivative of $\boldsymbol{c_t}(s,t)$ with respect to t is set equal to the left hand side of equation (3.12) as shown in equation (3.13).

$$\boldsymbol{c_t}(s,t) = \alpha \boldsymbol{c}''(s) - \beta \boldsymbol{c}''''(s) - \nabla \boldsymbol{E_{ext}} \tag{3.13}$$

The solution of equation (3.13) becomes steady when the term $c_t(s,t)$ approaches zero and reaches the objective of equation (3.12).

The GVF snake method introduced the GVF field as the static external force field that is derived from the image data and does not change as the snake deforms. The GVF field is defined as $\boldsymbol{F}_{ext} = \mathrm{V}(x,y)$. To obtain the corresponding dynamic snake equation, the potential force $-\nabla \boldsymbol{E_{ext}}$ in equation (3.13) is replaced with $V(x,y)$. The GVF field is given by the equation $V(x,y) = [u(x,y), v(x,y)]$, where u and v represent the direction and strength of the field, and can be solved by minimizing the energy function of the relation in equation (3.14).

$$\varepsilon = \int\int \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla \mathbf{f}|^2 |V(x,y) - \nabla \mathbf{f}|^2 dxdy \tag{3.14}$$

where μ is a regularisation parameter controlling the effects between the first term and the second term in the integrand. This parameter is set according to the amount of noise present in the image (more noise, increase μ). $\mathbf{f}$ is an image map. $\nabla \mathbf{f}$ is a gradient of an

edge map having vectors pointing toward the edges and normal to the edge ($\nabla \mathbf{f} = (\boldsymbol{f_x}, \boldsymbol{f_y})$) . $\boldsymbol{u}$ and $\boldsymbol{v}$ is initially set to ${\boldsymbol{f_x}}^2$ and ${\boldsymbol{f_y}}^2$ respectively.

The energy function of equation (3.14) can be solved by using calculus of variations (160) to find a solution for the Euler equations in equation (3.15) and (3.16).

$$\mu \nabla^2 \boldsymbol{u} - (\boldsymbol{u} - \boldsymbol{f_x})\left({\boldsymbol{f_x}}^2 + {\boldsymbol{f_y}}^2\right) = 0 \tag{3.15}$$

$$\mu \nabla^2 \boldsymbol{v} - \left(\boldsymbol{v} - \boldsymbol{f_y}\right)\left({\boldsymbol{f_x}}^2 + {\boldsymbol{f_y}}^2\right) = 0 \tag{3.16}$$

where ∇^2 is the Laplacian operator (a second order differential operator).

The value $\boldsymbol{u}$ and $\boldsymbol{v}$ can be solved by running the equations for t iterations as mathematically defined in equations (3.17) and (3.18).

$$u_t(x, y, t) = \mu \nabla^2 u(x, y, t) - (u(x, y, t) - f_x(x, y)) \cdot \left(f_x(x, y)^2 + f_y(x, y)^2\right) \tag{3.17}$$

$$v_t(x, y, t) = \mu \nabla^2 v(x, y, t) - (v(x, y, t) - f_y(x, y)) \cdot \left(f_x(x, y)^2 + f_y(x, y)^2\right) \tag{3.18}$$

3.3.2 Multiscale Normalised Cuts Segmentation

In a graph-based segmentation, a graph $G = (V, E, W)$ represents the relationship of nodes (vertices) V connected by edges E with weights W associated on each edge. In the MNcut, the graph G has the pixels as graph nodes V, and pixels within distance less than R are connected by a graph edge in E. The relationship of pixel connection is computed through a mixed weight value matrix $\boldsymbol{W_{Mixed}}$ (equation (3.21)), which measures the likelihood of pixels a and b belonging to the same coherent region. The $\boldsymbol{W_{Mixed}}$ is a combined two grouping cues, pairwise pixel affinities of intensity $\boldsymbol{W_I}$ (equation (3.19)) and pairwise pixel affinities of intervening contour $\boldsymbol{W_C}$ (equation (3.20)).

$$W_I(a,b) = e^{-\frac{\|X_a - X_b\|^2}{\sigma_x} - \frac{\|I_a - I_b\|^2}{\sigma_I}} \tag{3.19}$$

$$W_C(a,b) = e^{-max_{x \in line(a,b)} \frac{\|Edge(x)\|^2}{\sigma_C}} \tag{3.20}$$

$$W_{Mixed}(a,b) = \sqrt{W_I(a,b) \times W_C(a,b)} + \alpha W_C(a,b) \tag{3.21}$$

where X_a and I_a is the location and intensity of pixel a; σ_x, σ_I, and σ_C is a variance of pixel location, intensity, and edge; $line(a,b)$ is a straight line joining pixels a and b; $Edge(x)$ is the edge strength at location x; α is a parameter, contributing to the magnitude of $W_{Mixed}(a,b)$.

The bipartition of a graph $V = A \cup B$ is based on the Normalised Cuts algorithm, partitioning the graph by maximising the ratio of affinities with a group to that across groups, as defined in equation (3.22) and (3.23).

$$Ncut(A,B) = \frac{Cut(A,B)}{Volume(A) \times Volume(B)} \tag{3.22}$$

where

$$Cut(A,B) = \sum_{a \in A, b \in B} W(a,b) \tag{3.23}$$

The binary group vector is defined as $X_A \in \{0,1\}^N$, with $X_{A,a} = 1$ if pixel a belongs to segment A, otherwise 0 if the pixel belongs to segment B. The segmentation follows the criteria:

$$maximise \ \frac{1}{2} \sum_{l=1}^{2} \frac{\boldsymbol{X_l^T W X_l}}{\boldsymbol{X_l^T D X_l}} \tag{3.24}$$

where $\boldsymbol{D}$ is a diagonal matrix, $D(a,a) = \sum_j W(a,b)$, $\boldsymbol{W}$ is the mixed weight values of affinity graph.

In multiscale graph segmentation, the graphs are decomposed into different scales and are connected by three different types of matrices: the multiscale affinity matrix, the cross-scale interpolation matrix, the cross-scale constraint matrix. To connect between scales, pixels are interpolated by using the cross-scale interpolation matrices $\boldsymbol{C_{s,s+1}}$ defined in equation (3.25).

$$C_{s,s+1}(a,b) = \begin{cases} \frac{1}{|N_a|} & if\ b \in N_a \\ 0\ otherwise \end{cases} \tag{3.25}$$

where N_a is sampling neighbourhood of a.

To partition graph into different scales, the matrix $\boldsymbol{C}_{s,s+1}$ together with the cross-scale constraint matrices $\boldsymbol{C}$ are employed to control the relationship between nodes in layer $\boldsymbol{I_s}$ and $\boldsymbol{I_{s+1}}$. The matrix $\boldsymbol{C}$ has values in a diagonal direction as shown in equation (3.26).

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{C_{1,2}} & & -\boldsymbol{I_2} & & 0 \\ 0 & \ddots & \boldsymbol{C_{S-1,S}} & \ddots & -\boldsymbol{I_S} \end{pmatrix} \tag{3.26}$$

The cross-scale constraint equation satisfies the condition in equation (3.27).

$$\boldsymbol{CX} = 0 \tag{3.27}$$

where $\boldsymbol{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_s \end{pmatrix}$ and $X_s \in \{0,1\}^{N_s^* \times K}$ is the partitioning matrix at scale s, $N_s^* = \sum_s N_s$, $X_{s+1}(a) = \frac{1}{|N_a|}\sum_{b \in N_a} X_s(b)$. The neighbourhood N_i specifies the projection of $a \in \boldsymbol{I_{s+1}}$ on the finer layer $\boldsymbol{I_s}$.

The multiscale affinity matrix covers the full range of affinity matrix in each scale and is defined as the addition of affinity matrix from each scale as shown in equation (3.28).

$$\boldsymbol{W_{Full}} = \boldsymbol{W_1} + \boldsymbol{W_2} + \cdots + \boldsymbol{W_s} \tag{3.28}$$

This $\boldsymbol{W_{Full}}$ is memory inefficient and become unmanageable when an image size is increasing. As a result a compressed multiscale affinity matrix $\boldsymbol{W_{reconstruct}}$ is reconstructed by using cross-scale interpolation matrix $\boldsymbol{C_{s,s+1}}$ and recomputing $\boldsymbol{W_s}$ by either sub-sample image or sampling values from the affinity matrix $\boldsymbol{W_{s-1}}$. The reconstruction affinity matrix is computed by using equation (3.29).

$$\boldsymbol{W_{reconstruct}} = \boldsymbol{W_1} + \boldsymbol{C_{1,2}^T W_2^c C_{1,2}} + \cdots + \boldsymbol{C_{s,s+1}^T W_s^c C_{s,s+1}} \tag{3.29}$$

For the multiple partition, a generalized K-way Ncut function can be similarly defined using $\boldsymbol{X} = [X_1, \ldots X_K]$. The optimal multiscale normalised cut partitioning can be solved by computing the K eigenvectors, corresponding to the K largest eigenvalues, with the maximising criteria of equation (3.30). The K largest eigenvalues is determined by the number of segments altered by a user. Selection of K values depends on the complexity of tumour structure. The higher the number of K values, the finer the composition is acquired. It is obvious that the higher the K values, the longer the computation time.

$$maximise\ \frac{1}{K}\sum_{l=1}^{K}\frac{\boldsymbol{X_l^T W X_l}}{\boldsymbol{X_l^T D X_l}} \tag{3.30}$$

Graphically, one-dimensional view of multiple-scale graph decomposition with R = 1 is decomposed into three scales, $s \in \{1, 2, 3\}$, as shown in Figure 3.4. At each scale, pixels are connected by the relationship defined by the affinity matrices: $\boldsymbol{W_1}$, $\boldsymbol{W_2}$ and $\boldsymbol{W_3}$ for scales 1, 2 and 3 respectively. The connections between scales are defined by the cross-scale interpolation matrices: $\boldsymbol{C_{1,2}}$ and $\boldsymbol{C_{2,3}}$. Pixels at each scale are sampled at $(2R + 1)^{s-1}$. The value R determines the relationship of two pixels in a graph, i.e. the two pixels are connected if they are within distance R. The length of R offers distinct segmentation results. Generally, a larger R produces better segmentation because a long range graph

connections facilitate propagation of local grouping cues across larger image regions. However, the larger R requires longer segmentation time.

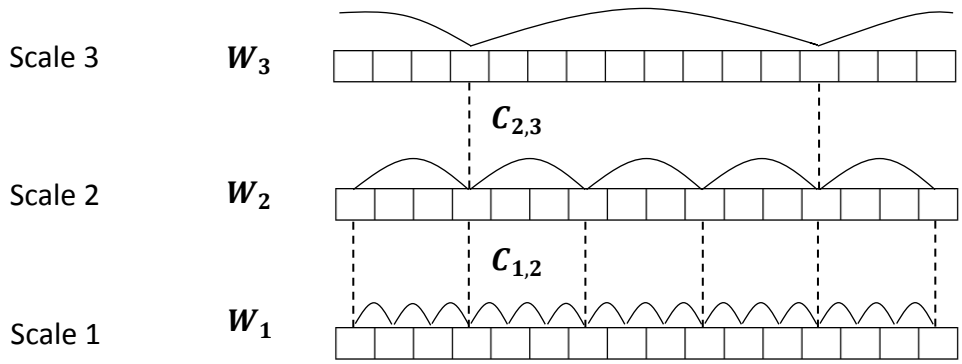


Figure 3.4: Three scales of the MNcut graph decomposition with R = 1. Relationship between pixels at each scale and cross scales are defined by affinity matrices W_s and cross-scale interpolation matrices $C_{s,s+1}$.

3.3.3 Evaluation of Segmentation

An evaluation procedure is important to report the reliability and quality of segmentation techniques. Evaluation methods for image segmentation can be divided into two main approaches: subjective evaluation and objective evaluation (161). Subjective evaluation is a commonly used approach which requires one or more human experts to visually compare the image segmentation results, derived from different algorithms or image types. Objective evaluation approaches can be subdivided into supervised or unsupervised. Supervised evaluation approaches compare machine segmented images against manually segmented images or reference images (also called gold standard or ground truth) (139, 162). Unsupervised evaluation approaches do not require human visual assessment or manual segmentation. Instead, it relies on automatic validation by utilising different segmentation techniques or different parameterisation from a specific segmentation method to quantify the quality of segments. Compared with the

unsupervised evaluation technique, the supervised evaluation approach is the preferred method in this thesis because of the complexity of brain tissue composing various components in the context of childhood brain tumours, such as abnormal tissues like solid tumour, cyst, necrosis, and normal tissues like white matter, grey matter, and CSF (9).

In the literature, several studies applying brain tumour segmentation as part of their analysis have been conducted on structural MRI, for example on T1-weighted images (162, 163), FLAIR images (164, 165) and T2-weighted images (166, 167). A combination of parameters derived from multimodal MRI has also been applied to perform posterior fossa tumour segmentation (168-170). Therefore, it is observed that brain tumour segmentation has been broadly studied based on both single modality and multimodal MRI. However, a segmentation method has not yet been applied to different types of MR image-based paediatric brain tumours in arbitrary brain location.

As aforementioned review, the quality of the proposed segmentation method is measured by using two evaluation methods: a scoring system by human experts and a supervised evaluation. A scoring system method adapts a five-point Likert scale to quantify the quality of the method performed on four types of MRI (T2, FLAIR, pre-contrast T1 and post-contrast T1) to identify which MRI sequence provides the better tumour outlines.

The supervised evaluation quantifies the degree of similarity between tumour volumes obtained from the manual segmentation and the proposed semi-automatic segmentation method. The common metrics has been reported are sensitivity, specificity, Hausdorff distance, dice coefficient and volume overlapping ratio, adapted in this thesis.

3.3.3.1 Intraclass Correlation Coefficient

To indicate the degree of reliability of rating and compare the segmentation outcome from four types of MRI, the inter-rater reliability is calculated. The inter-rater reliability measures the level of which the ratings from different evaluators are the same when expressed as deviation from their means. A high inter-rater reliability indicates the high similarity of rating scores of each item provided by all evaluators, although the absolute values of rating score may differ from one to another evaluator (171). Therefore, a high inter-rater reliability is preferable as this would mean that the segmentation results are rated consistently or the manual segmentation is consistent across evaluators. A low inter-rater reliability could imply that the segmentation results are defective or the evaluators may need to be re-trained. In order to assess this inter-reliability, the intraclass correlation coefficient (ICC) approach proposed by Shrout and Fleiss (172) is adapted.

There are several forms of ICC which are implemented based on three main guidelines. First is that the study requires a one-way or two-way analysis of variance (ANOVA). Second is whether the judges' mean rating is relevant to the reliability of interest. Third is that an individual rating or the mean of several ratings is relevant. All of the ICC forms are defined as "a ratio of the variance of interest over the sum of variance of interest plus error". The ICC value varies between 0 and 1, indicating from no reliability to perfect reliability.

3.3.3.2 Sensitivity and Specificity

In brain tumour segmentation, the sensitivity indicates the tumour portion that the segmentation method correctly segmented, whereas the specificity indicates the non-

tumour portion that the segmentation method correctly segmented. The sensitivity and specificity are defined in equation (3.31) and (3.32) respectively.

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.31}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3.32}$$

where TP is true positive (correctly segmented as a tumour pixel), FN is false negative (incorrectly segmented as a non-tumour pixel), TN is true negative (correctly segmented as a non-tumour pixel), FP is false positive (incorrectly segmented as a tumour pixel).

3.3.3.3 Modified Hausdorff Distance

The Hausdorff distance (173) quantifies the degree of dissimilarity between two objects, with lower value indicating higher similarity between two objects and zero indicating a perfect match. The definition of modified Hausdorff distance (MHD) presented by Dubuisson and Jain (174) is illustrated as follows.

Given two finite point sets $A = \{a_1, a_2, \dots, a_m\}$ and $B = \{b_1, b_2, \dots, b_n\}$, the Hausdorff distance defines a distance between a point in A and B as the maximum of $h(A, B)$ and $h(B, A)$ as shown in equation (3.33).

$$H(A, B) = \max(h(A, B), h(B, A)) \tag{3.33}$$

The modified Hausdorff distance (MDH) defined $h(A, B)$ as shown in equation (3.34).

$$h(A, B) = \frac{1}{m} \sum_{a \in A, b \in B} min \, \|a - b\| \tag{3.34}$$

where $\|a - b\|$ is the Euclidean distance between two points $a \in A$ and $b \in B$.

3.3.3.4 Overlap Volume Ratios

The two overlap volume ratios and dice coefficient are used to measure the similarity between the semi-automatically and manually segmented volumes. Three measures are used in evaluation of the brain tumour segmentation, and are defined as follows:

$$Ratio1 = (V_1 \cap V_2)/ V_1 \tag{3.35}$$

$$Ratio2 = (V_1 \cap V_2)/ V_2 \tag{3.36}$$

$$Dice\ Coefficient = 2(V_1 \cap V_2)/(V_1 + V_2) \tag{3.37}$$

where V_1 is the semi-automatically segmented volume and V_2 is the manually segmented volume.

3.4 Texture Analysis

Tissue characterisation by quantitative MRI has revealed that malignant intracranial tumours have a propensity of longer spin-lattice relaxation times than benign intracranial tumours and normal brain tissue (175, 176). However, because of tissue heterogeneity, the relaxation times of high-grade and low-grade lesions overlap and therefore reduce accuracy of diagnosis (24, 29). Although an optimised setting of image acquisition was set, the performance of tissue characterisation was not noticeably improved (177). Texture analysis offers an alternative quantitative diagnostic tool for MR image analysis. TA has been used to assess MR images of biological tissues, which contain large amounts of microscopic details that are scarcely legible by visual inspection. In comparison with human visual perception, TA can better characterise patterns of tissues in MR images, because it is more sensitive to variations of grey-level intensity in image pixels. This has

been experimentally demonstrated by Herlidou et al. where TA was reported to have produced better discrimination between healthy and pathological tissues compared to visual examination (20). In addition, TA can be conducted automatically and therefore is independent of operator's expertise. In this thesis, TA is adapted to generate textural features from different types of MR images in order to support diagnostic classification of childhood brain tumours. The following sections, existing applications of TA on structural, diffusion and multimodal MRI are discussed.

3.4.1 Texture Analysis for the Characterisation of Childhood Brain Tumours

3.4.1.1 Texture Analysis of Conventional MRI

Conventional MRI offers structural imaging of anatomical elements and is a more preferable modality to functional imaging for localising tumour site, delineating tumour boundary and estimating tumour size (9, 178). Various studies have investigated texture analysis based structural MRI with T2 being the most informative for identifying pathology (24, 26, 43, 177, 179, 180). T2-weighted images can provide a rough measurement of free water abundance, macromolecules, and the presence of citrate (181). The sensitivity to such molecules enables T2-weighted images to provide tumour localisation with higher sensitivity and specificity than other MR images (164). It also allows T2-weighted images to be more sensitive in detecting and differentiating intracranial lesions than T1-weighted and FLAIR images (32, 164, 182, 183).

In studies of brain pathology, signal intensity and relative signal intensity of T2-weighted images have revealed identity of intracranial lesions. Arai et al. studied T2 signal intensity of pilocytic astrocytoma and medulloblastoma in children (82). They reported that the solid portion of cerebellar pilocytic astrocytoma is isointense to CSF on T2-weighted images and such signal intensity is not observed in medulloblastoma. Forbes et al. used the relative T2 signal intensity (T2 signal intensity tumours/T2 signal intensity vitreous tumour) to differentiate the three common paediatric brain tumour types (179). The relative T2 signal intensity identifies pilocytic astrocytoma as hyperintense, while ependymoma and medulloblastoma as hypointense. The signal intensity and relative signal intensity approaches demonstrate high accuracy in differentiating medulloblastomas and pilocytic astrocytomas. However, separating ependymoma from the two tumour types is still an unresolved problem.

In a more extensive TA study, Kjaer et al. reported that no discrimination between adult benign and malignant tumour growth was possible, with an analysis of the common first-order and second-order grey-level statistics on T1- and T2-weighted images (26). In a study conducted by Herlidou-Même et al. texture analysis of a multi-centre dataset of T1- and T2-weighted images for the characterisation of healthy and pathologic brain tissue (meningioma, gliomas and lymphoma) was investigated in adult subjects (24). The study reported that TA methods, including histogram, GLCMs, gradient matrix and GLRLMs, can moderately separate benign and malignant tumours based on T1- and T2-weighted images.

Recently, two studies demonstrated the use of TA applied on MR images of paediatric posterior fossa tumours with promising outcome. Fetit et al. employed the MaZda tool for dichotomous classification of three paediatric brain tumour types: embryonal against astrocytic, embryonal against ependymal and ependymal against astrocytic (184). Even sized square region of interests were manually segmented. Cystic or other abnormal tissue regions were not mentioned to be excluded. Textural features were derived from histogram, absolute gradient, GLCM, GLRLM, wavelets and autoregressive model. Their preliminary study showed that textural features extracted from T2-weighted images produced promising results, particularly the classification of embryonal against astrocytic. However, their results showed lower accuracy figures of separating ependymal tumours from embryonal and astrocytic tumour types.

Vlachou et al. also employed the MaZda software to classify paediatric posterior fossa tumours based on textural features of pre-contrast T1- and T2-weighted images (19). Region of interests excluding large cystic areas were manually segmented. Their study demonstrated high sensitivity and specificity of medulloblastoma, pilocytic astrocytoma and ependymoma. However, sensitivity of classifying ependymoma is relatively lower than that of classifying the other two tumour types. In terms of texture analysis, GLCM-based textural features were mentioned to give higher discriminative power than textural features derived from other methods.

As discussed above, MRI, especially T2-weighted images, are useful for revealing structural information and sensing intracranial lesions. In terms of TA-based classification, some studies have reported poor to moderate differentiation between benign and

malignant tumour growth with the use of common statistical texture analysis on T1- and T2-weighted images (24-26). However, some studies demonstrated promising results of diagnostic classification of posterior fossa tumours in children with the use of a larger set of textural features calculated from statistical analysis and wavelet transform methods.

Despite promising results, separating ependymomas from medulloblastomas and pilocytic astrocytomas in some patients are still challenging. Additional non-invasive diagnostic information is still needed; consequently, the textural features of diffusion MRI are examined in our study and discussed in the following section.

3.4.1.2 Texture Analysis of Diffusion MRI

Tissue microstructure can be reviewed by using diffusion MRI, which offers two main functional imaging techniques: DWI and DTI. Both techniques provide measurement of water movement; however the rate of movement is reported in different patterns. DWI reflects the rate of water diffusion in tissues described by a reconstructed parameter ADC, whereas DTI probes diffusion anisotropy in tissues characterised by a reconstructed parameter FA.

Diffusion anisotropy in white matter measured by DTI has potential assessment of brain functionality and diseases, such as brain maturation in children, newborns or premature babies and diagnosis of white matter disorder in children (135). Although DTI has also been studied for characterising childhood brain tumours (185), DWI is more routinely acquired and is popularly employed for the study of brain related pathology.

DWI has found to be useful for distinguishing necrosis from cyst formation or oedema, identifying different tumour types and delineating pathological tissue boundaries against normal cerebral tissues (28). DWI was also reported to provide information that complements conventional pre- and post-contrast structural imaging (186). An ADC map, a reconstructed image obtained from DWI can reflect the rate of microscopic water diffusion better than DWI at each b-value because it considers signal at two different b-values (57).

ADC has been used as a biomarker in neurological studies including classification of brain tumours (187) and evaluation of brain tumour infiltration (188). Average of ADC values, tumour ADC ranges, or relative ADC ratios are used as measurement for tumour discrimination. Some studies employed ADC values in characterising brain pathology are described as follows.

Kono et al. showed that ADC values can predict the degree of malignancy of adult astrocytic tumours (189). However there is an overlap between ADC values of grade II astrocytomas and glioblastomas.

Bulakbasi et al. presented that ADC values and ratios can provide higher sensitivity in differentiating adult low-grade and high-grade brain tumours, than conventional contrast-enhance MRI (190).

Yamasaki et al. evaluated average intensity of ADC maps in children and adults (191). Their study showed a promising discrimination outcome among various tumour groups by using logistic discriminant analysis with ADC, age, and sex being independent variables.

Oksuzler et al. studied the relationship of diffusion properties in normal children and in different paediatric cerebral diseases (183). They demonstrated that ADC values are associated with paediatric neurologic disorders.

Rumboldt et al. studied the classification of four childhood brain tumour types: juvenile pilocytic astrocytoma, ependymoma, medulloblastoma and atypical teratoid/rhabdoid tumours (192). They reported the usefulness of ADC maps with the use of absolute ADC values of contrast-enhancing solid tumour regions and ADC ratios (ADC of solid tumours to ADC of normal-appearing white matter). Their derived ratios corresponded with values of previous studies presented by Le Bihan (185) and Yamasaki et al. (191). The corresponding figures showed that ADC values of juvenile pilocytic astrocytoma, ependymoma, medulloblastoma and atypical teratoid/rhabdoid tumours ranges from high to low intensity respectively.

Tozer et al. showed that ADC histograms may predict untreated low-grade glioma subtypes; astrocytomas and oligodendrogliomas in adult patients (193). Region of interest excluded large cystic region and biopsy damage.

Jaremko et al. used minimum ADC values to characterise three main paediatric posterior fossa tumour types: juvenile pilocytic astrocytoma, medulloblastoma and ependymoma (194). They suggested that cystic, necrotic and haemorrhagic areas must be avoided and the region of interest should be within the solid tumour.

Bull et al. used histogram analysis of ADC values to classify the common posterior fossa tumours of paediatric patients (195). They selected tumour region where cyst having a

size greater than solid tumour portion is excluded. Their study showed high classification accuracy with the use of logistic regression method.

Rodriguez Gutierrez et al. also used ADC-based histogram analysis to differentiate single-centre dataset of paediatric posterior fossa tumours (30). Combining 25th and 75th percentile and skewness features derived from ADC-based histogram analysis was reported to achieve high classification accuracy and outperformed classification of post-contrast T1 and T2 based features.

Several studies have shown that ADC-based histogram analysis can differentiate paediatric posterior tumours with high classification accuracy. However, multi-centre dataset of ADC maps with extensive TA has not been reviewed. In addition, although an optimal region of interest is selected, ADC values of medulloblastomas, pilocytic astrocytomas and ependymomas still overlap. In particular, ADC values of ependymomas are typically between and overlap with medulloblastomas and pilocytic astrocytomas. This overlapping could be contributed by technical difficulties in ADC measurement (small cystic, haemorrhagic or calcific components) and variation in tumour pathology.

ADC in combination with MRI can provide an enhanced discrimination of brain tumour types (188, 194). As a result, it is hypothesised that the combination of conventional and diffusion MRI based TA could provide more comprehensive information for the characterisation of childhood brain tumours. The review of multimodal MRI based TA is discussed in the following section.

3.4.1.3 Texture Analysis of Multimodal MRI

Diffusion MRI has been reported to provide better information than structural MRI for discriminating histological types of brain tumours. However, analysing diffusion MRI alone still produces overlapping feature values among the three tumour types. Various techniques of multimodal MRI analysis have shown higher classification accuracy to differentiate tissue pathology than using a single imaging modality, described as follows.

Bulakbasi et al. showed that combined metabolites ratio acquired from single-voxel proton MR spectroscopy (^{1}H-MRS) and ADC values better differentiates and grades adult brain tumours than using MRI or using single technique alone (196).

Tzika et al. studied the relationships between a number of parameters; MSRI, hemodynamic MRI (HMRI) and ADC values in order to examine physiologic state of tumours in children (197). They suggested that multi-parametric MRI with such parameters can provide an enhanced assessment of paediatric brain tumours.

Costanzo et al. assessed multi-parametric MR of gliomas based on ^{1}H-MRSI, perfusion weighted imaging (PWI) and DWI to discriminate infiltrating tumour from surrounding vasogenic oedema or normal tissues, and high- from low- grade gliomas in adult patients (198). They suggested that this approach may provide useful information for guiding stereotactic biopsies, surgical resection and radiation treatment.

Schneider et al. studied the multi-parametric analysis of combined ADC values and ^{1}H-MRS in differentiating posterior fossa tumours: medulloblastomas, ependymomas, pilocytic astrocytomas and infiltrating gliomas in a paediatric population (199). They

showed that the multi-parametric analysis provides promising discrimination among these tumour types.

Verma et al. created an intensity-based tissue profile derived from multi-parametric intensity of B0, DWI, FLAIR, T1, and gadolinium-enhanced T1, ADC and FA (200). Their preliminary results demonstrated that multi-parametric tissue characterisation helps to better differentiate among neoplasm, oedema, and healthy tissue in adult patients.

Forbes et al. presented combining DWI and T2 to classify medulloblastoma, ependymoma and pilocytic astrocytoma (84). The relative intensity (tumour/ white matter) of DWI and T2 were used as features in a decision tree classification to determine the likelihood of paediatric cerebellar neoplasms. Medulloblastoma is likely DWI hyperintense. Ependymoma is DWI hypointense and T2 hypointense. Pilocytic astrocytoma is likely DWI hypointense and T2 hyperintense.

From the aforementioned review, several studies have investigated features from various types of MR-based signal and image acquisition (e.g. HMRS/ MRSI, PWI, HMRI, MRI) combined with ADC or DWI values. Most studies applied first-order statistical TA on multimodal MRI, but second- and higher-order statistical analysis and wavelet transform have not been extensively explored. In addition, as summarised in Table 3.1, multimodal MRI based TA has been reviewed to provide better classification of brain tumour types than conventional and diffusion MRI based TA. Therefore the integration of first-, second- and higher-order statistics and wavelet transform based texture derived from conventional and diffusion MRI is investigated in this thesis. This investigation aims to develop a more accurate and robust classifier for childhood brain tumours in a multi-

centre setting. In the following sections, texture analysis methods are reviewed in order to choose suitable methods for this thesis.

Table 3.1: The common texture analysis methods applied on conventional, diffusion and multimodal MRI of brain tumours.

Conventional MRI	Diffusion MRI	Multimodal MRI
• T2 images are more sensitive to pathology and provide better textural features for characterising brain tumours, compared to T1 and FLAIR images. • The most common TA methods found are statistical analysis and wavelet transforms which revealed promising discrimination of posterior fossa tumours.	• ADC maps are extensively studied with first-order statistical analysis being the most commonly applied. • A large combination of statistical and wavelet transform based textural features derived from ADC maps has not been comprehensively demonstrated.	• ADC values combined with values from other modality offer higher diagnostic accuracy than using single modality. • Statistical and wavelet transform based textural features of a combined ADC and MRI are not found in the literature.

3.4.2 Review of Texture Analysis Methods

Texture exhibits visual patterns which are properties of surfaces containing information about the structural arrangement. In the context of medical image processing, texture is used to describe tissue structures and functionality with terms such as smoothness, granulation, regularity, homogeneity, heterogeneity, etc. These parameters can be derived from local intensity variation by using texture matrices (201). Such matrices can be calculated by considering for example “grey-level patterns, pixel interrelationships and spectral properties of an image” within a region of interest (18).

Texture analysis was firstly introduced by Haralick et al. (202) and was applied in diversified applications from macroscopic satellite images (203) and material structures (204) to microscopic images of biological tissues (205). TA methods are broadly categorised into four main approaches namely structural, model-based, statistical and transform-based techniques (206, 207).

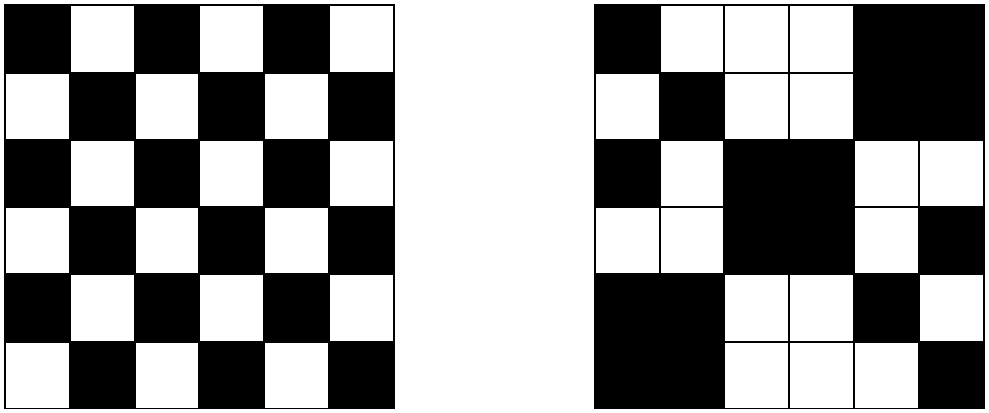
Structural-based methods, firstly presented by Haralick (205), symbolize texture by applying the principle of mathematical morphology to define a primitive (microtexture) to be placed at a particular location (spatial arrangement or macrotexture). These methods are more useful for producing new objects, by using approaches such as addition and subtraction between elements than quantifying surface arrangement. The natural textures of brain tumours, having variability of microstructure, macrostructure and limited information of separation lines between them, can be poorly defined by these methods (208).

Model-based methods represent image texture-based on mathematical models, such as fractal or stochastic method (209, 210). The fractal model (211) can be useful for discriminating some natural textures. However, it is less flexible to provide orientation selectivity and is not appropriate to describe local image structures (206). The stochastic model requires the computational complexity to estimate texture parameters.

Statistical techniques analyse the spatial distribution of images' grey-levels. The methods can be classified as first-, second- or higher-order statistics according to the number of pixels used to define the local feature. In first-order statistics, local features are solely obtained from individual pixel values. Second-order statistics construct matrices based on

pixel pair relationships and higher-order statistics consider number of consecutive pixels at each grey-level.

The first-order statistics include mean, variance, skewness and the kurtosis. However, some of these parameters can be roughly estimated visually, while provide limited information about the relative position of pixels to each other. For example, a checkerboard pattern in Figure 3.5 (a) has an equal number of black and white pixels, and produces the same grey-level histogram as Figure 3.5 (b).



(a) (b)

Figure 3.5: Two images having the same textural features of first-order statistics. Image (a) and (b) have the number of black and white elements.

As a result, it is difficult to differentiate the two images using the first-order statistical analysis (212). The second- and higher-order statistical methods, constructed from grey-level co-occurrence matrices (GLCMs) and grey-level run-length matrices (GLRLMs) respectively, have a tendency to obtain higher discrimination indices than the first-order statistics and visual examination (23). Although GLCMs and GLRLMs can describe surface pattern better, their matrices are rotational-variant (213), i.e. the rotation of the same

object produces different texture matrices and can lead to misinterpretation of object's pattern.

Transform-based TA methods employ time-frequency analysis based on Fourier (214), Gabor (215) or Wavelet transforms (216). These methods transform an image to another space, using particular mathematical functions. The texture characteristics of a transformed image space are commonly represented by energy and frequency. Fourier transform has the potential to capture an image's global features, however, it can be insufficient to define local features (217) and to provide spatial localisation (206). Imaging features of intracranial lesions that vary in space and time may not be captured by the Fourier transform method. Gabor filters offer an improved spatial localisation, however, their non-orthogonality application produces redundant features at different scales and is difficult to localise a spatial structure of natural textures (218, 219). Wavelets are an extension of windowed Fourier analysis (220), decomposing an image into multi-scale of spatial resolutions. Wavelets provide several advantages over the other two transform approaches. For example, a multiresolution decomposition can represent natural texture at a suitable scale with lower computational cost than Gabor filters (216); an orthogonality property of certain wavelet families produces non-redundant feature maps from an original image to different scales. In spite of these advantages, wavelet transforms are translation-variant (221) which can lead to translation-variant textural features, reducing classification accuracy. However, they can possibly supply additional textural properties that may not be captured in GLCMs and GLRLMs. In addition, wavelet transform methods are often found in the medical image analysis literature, including image-based TA software, such as MaZda tool (222).

The statistical methods potentially provide better classification outcome than the structural or transform based methods (208, 223). In TA-based brain studies, the first-order statistics, GLCMs, absolute gradient matrices, GLRLMs and wavelets are the well-known techniques for the characterisation of healthy and pathological human cerebral tissues (23, 24, 224). Among these approaches, the first- and second-order statistical analysis outperform wavelet transforms in the classification of Alzheimer's disease based on T2-weighted brain images (225), and the most popular texture analysis methods for MR images was voted to be the GLCMs (23). In addition, compared with features derived from Laws and Haar wavelet, GLCM-based features can achieve higher accuracy in differentiating between anaplastic/large cell and non-anaplastic/large cell medulloblastoma in children (226). Dange et al. compared texture classification based on four methods: GLCMs, Haar, Daubechies-4, Symlet-8 (213). The study showed that in the case of rotation, Haar wavelet was the most efficient method in terms of classification accuracy and computation time. Although GLCMs can provide relative similar classification accuracy, it performed badly when images are rotated.

As discussed above, GLCMs are generally preferable in several applications. However, GLCMs require higher computation time and yield less accuracy when rotation takes place. Wavelet transform based features has been reported a high classification rate of some types of brain tumours (227). Therefore, wavelet transform based features are recommended as complementary features and offer superior result when image rotation is considered. In order to summarise the discussion aforementioned, the advantages and disadvantages of TA methods for diagnosing brain tumours are shown in Table 3.2.

Table 3.2: Advantages and disadvantage of TA for defining brain tumours.

Textural Analysis	Advantages	Disadvantages
Structural method e.g. morphological processing	It is useful for representing an object by using addition and subtraction between elements.	Micro and macrostructure of brain tumours' MR images and faint separation lines between them can be poorly defined.
Model-based method e.g. Fractal and Stochastic method	It offers sophisticated models for defining brain tumours.	It requires high computation time and provides deficiency of orientation selectivity.
Statistical analysis e.g. first-order statistics, GLCM, GLRLM	The first-order statistic is simply calculated and can describe simple structures of biological tissues. The second- and higher-order statistics provide higher discriminative power and are found to be useful in various studies.	The first-order statistic can give the same textural features of two distinct objects having the same grey-level distribution. The second- and higher-order statistics require high computation time and shift-variant.
Transform-based method e.g. Fourier, Gabor and wavelet transforms	Wavelet transforms offer multiscale decomposition which can provide a suitable scale for brain tumours' MR images with low computational cost.	Discrete wavelet transform is shift-variant and energy derived from wavelet coefficients can be shift-variant, resulting in low classification accuracy.

As critically reviewed, this thesis aims to combine the advantages of the most commonly used sets of textural features derived from statistical and wavelet transform approaches. The first-, second- and higher-order statistics as well as Haar, Daubechies and Symlet wavelets are selected for our study. The theoretical background of these texture analysis methods and their associated textural features are explained in the following sections and Appendix A.

3.4.3 First-Order Statistics Based Textural Features

The first-order statistical based textural features quantify an average brightness and an intensity distribution within a region of interest. Let x_i be the grey-level intensity and n_p be the number of pixels within the ROI. Mean, variance, skewness and kurtosis are defined in equation (3.38) to (3.41) respectively.

Mean
$$m = \frac{1}{n_p} . \sum_{i=1}^{n_p} x_i \tag{3.38}$$

Mean is the average grey-level intensity within the ROI.

Variance
$$var = \frac{1}{n_p} . \sum_{i=1}^{n_p} (x_i - m)^2 \tag{3.39}$$

Variance describes the deviation of the intensity from the mean within the ROI.

Skewness
$$sk = \frac{\frac{1}{n_p} \sum_{i=1}^{n_p} (x_i - m)^3}{\left(\frac{1}{n_p} \sum_{i=1}^{n_p} (x_i - m)^2\right)^{3/2}} \tag{3.40}$$

Skewness represents the degree of data asymmetry around the mean. Value of zero indicates the symmetrical distribution around the mean. Positive or negative values indicate the asymmetric distribution with a long tail to the right and with a long tail to the left respectively.

Kurtosis
$$k = \frac{\frac{1}{n_p} \sum_{i=1}^{n_p} (x_i - m)^4}{\left(\frac{1}{n_p} \sum_{i=1}^{n_p} (x_i - m)^2\right)^{2}} - 3 \tag{3.41}$$

Kurtosis measures the peak of data distribution compared with the Gaussian shape. Kurtosis is zero if the data distribution has a Gaussian distribution, otherwise negative

or positive if the distribution is flatter or has higher peak than a Gaussian distribution respectively.

3.4.4 Absolute Gradient Matrix Based Textural Features

Absolute gradient measures an intensity change across each pixel. Pixels laying on a boundary of an image are not considered. An example absolute gradient value of a pixel **M** in an image of Figure 3.6 is calculated by using the three neighbouring pixels as shown in equation 3.42. Textural features derived from the absolute gradient matrix are the same as the first-order statistical features.

A	B	C	D	E
F	G	H	I	J
K	L	**M**	N	O
P	Q	R	S	T
U	V	W	X	Y

Figure 3.6: An example image for calculating an absolute gradient matrix. The capital letters represent grey-level intensity of pixels.

For 3 neighbouring pixels:

$$\boldsymbol{A} = \sqrt{(H-R)^2 + (N-L)^2} \tag{3.42}$$

3.4.5 Grey-Level Co-Occurrence Matrix Based Textural Features

Haralick et al. introduced the GLCM, a second-order statistical analysis (202). GLCMs are constructed by considering the relationship between pixel pairs across the entire intensity range and tabulating the frequency of grey levels. The GLCMs have the row and column which is equal to the number of grey levels. GLCMs count the number of pairs of grey tone pixels at a distance d within a ROI in four directions (θ): 0°, 45°, 90^0, and 135^0.

The GLCMs can be reconstructed as shown in Figure 3.7. Given an example image (Figure 3.7 (a)) with five grey-levels (0, 1, 2, 3, and 4); the GLCMs with d = 1, scanned in four directions, are constructed as shown in Figure 3.7 (b) - 3.7 (e). The GLCMs show different patterns, for example the pixel pair (2,4) is counted differently in the four GLCMs. In the horizontal (0°), the diagonal (45°), the vertical (90°) and the anti-diagonal (135°) direction, the pixels pair (2,4) appears 4, 1, 1 and 0 times respectively. In this thesis, the GLCMs are constructed with d = 1, 2, 3, 4 and 5; θ = 0°, 45°, 90^0, and 135^0 for each distance d. The 20 textural features derived from each GLCM are used in this thesis (Table 3.3). The details of these textural features are described in section A.1, Appendix A. Theses textural features are calculated by using the built-in function in MATLAB and a package implemented by Avinash Uppuluri, available at File Exchange - MATLAB Central.

Although, GLCM approach can provide the spatial interrelationships of the grey tones and is invariant under monotonic grey tone transformations, there are several short-comings of this technique. Firstly, GLCM does not capture the shape features of the tonal primitives (tonal primitives are region of an image which can be evaluated in terms of shape and area, for example), and so it is less possibly to provide good information of large area primitives (205). Secondly, GLCMs requires high computation time because of generation of large matrices with different distance and degree in order to define texture structure. Thirdly, as the GLCMs can be very large, they may contain larger amount of data than the original image (228).

Despites the aforementioned drawbacks, Ying et al. showed the usefulness of GLCM-based textural features in separating subtype of medulloblastoma in children (226).

However, in order to obtain more insightful information of pathological tissue, grey-level run-length matrices are introduced to represent information that cannot be captured from GLCMs.

0	0	0	0	0
4	1	2	4	4
3	1	3	2	4
3	1	2	4	3
4	2	4	3	3

(a) Digital image

Grey-level

Grey-level	***0°***	***0***	***1***	***2***	***3***	***4***
0	4	0	0	0	0
1	0	0	2	1	0
2	0	0	0	0	4
3	0	2	1	1	0
4	0	1	1	2	1

(b) GLCM at d = 1 and θ = 0°

Grey-level

Grey-level	***45°***	***0***	***1***	***2***	***3***	***4***
0	0	0	0	0	0
1	1	0	1	1	0
2	1	0	2	0	1
3	0	2	0	1	1
4	2	1	0	0	2

(c) GLCM at d = 1 and θ = 45°

Grey-level

Grey-level	***90°***	***0***	***1***	***2***	***3***	***4***
0	0	0	0	0	0
1	1	2	0	0	0
2	1	1	0	1	1
3	0	0	1	2	3
4	3	0	2	1	1

(d) GLCM at d = 1 and θ = 90°

Grey-level

Grey-level	***135°***	***0***	***1***	***2***	***3***	***4***
0	0	0	0	0	0
1	1	0	0	1	1
2	1	1	1	1	0
3	0	1	2	0	1
4	2	1	0	1	1

(e) GLCM at d = 1 and θ = 135°

Figure 3.7: The four grey-level co-occurrence matrices of a digital image are computed with one pixel distance and scanned at 0°, 45°, 90° and 135°.

Table 3.3: List of 20 textural features derived from GLCM. The definitions of these textural features are described in Appendix A.

Textural feature	Textural feature
Autocorrelation ($autoc$)	Sum of squares ($sosv$)
Correlation ($corr$)	Sum average ($savg$)
Contrast ($contr$)	Sum variance ($svar$)
Cluster prominence ($cprom$)	Sum entropy ($sent$)
Cluster shade ($cshad$)	Difference variance ($dvar$)
Dissimilarity ($dissi$)	Difference entropy ($dent$)
Energy ($ener$)	Information measure of correlation 1 ($imc1$)
Entropy ($entro$)	Information measure of correlation 2 ($imc2$)
Homogeneity ($homo$)	Inverse difference normalised (idn)
Maximum probability ($maxpr$)	Inverse difference ($idmn$)

3.4.6 Grey-Level Run-Length Matrix Based Textural Features

Galloway (229) firstly proposed GLRLMs, which are categorised as higher-order statistical analysis. GLRLMs are constructed by systematically considering a number of consecutive pixels having the same grey-level at each scanning direction. This application was not recommended by Weszka et al. (223) and Conners and Harlow (230) because of its inefficiency compared with other traditional features, such as GLCM and power spectrum based features. Xiaoou (231), Chu et al. (232) and Dasarathy and Holder (233) described a new feature set of GLRLMs which improved classification accuracy over the traditional method.

Considering an example image with five grey-levels (0, 1, 2, 3 and 4) in Figure 3.8 (a), the GLRLMs can be computed at four directions: 0°, 45°, 90^{0}, and 135^{0} (Figure 3.8 (b) - (e)), showing different number of runs at each grey-level and each direction. For example, the number of run of grey-level 0 is counted as 5 horizontally (Figure 3.8 (b)) and 0 for other directions (Figure 3.8 (c) - (e)).

The 11 textual features derived from each GLRLM are used in this thesis (Table 3.4). The details of the GLRLM-based textural features are described in section A.2, Appendix A. These GLRLMs and their textural features are computed by using a package implemented by Xunkai Wei, available at File Exchange – MATLAB Central.

0	0	0	0	0
4	1	2	4	4
3	1	3	2	4
3	1	2	4	3
4	2	4	3	3

(a) Digital image

Number of run

Grey-level	***0°***	***1***	***2***	***3***	***4***	***5***
0	0	0	0	0	1
1	3	0	0	0	0
2	4	0	0	0	0
3	4	1	0	0	0
4	5	1	0	0	0

(b) GLRLM at $\theta = 0^{o}$

Number of run

Grey-level	***45°***	***1***	***2***	***3***	***4***	***5***
0	5	0	0	0	0
1	3	0	0	0	0
2	1	0	1	0	0
3	4	1	0	0	0
4	4	0	1	0	0

(c) GLRLM at $\theta = 45^{o}$

Number of Run

Grey-level	***90°***	***1***	***2***	***3***	***4***	***5***
0	5	0	0	0	0
1	0	0	1	0	0
2	4	0	0	0	0
3	2	2	0	0	0
4	5	1	0	0	0

(d) GLRLM at $\theta = 90^{o}$

Number of Run

Grey-level	***135°***	***1***	***2***	***3***	***4***	***5***
0	5	0	0	0	0
1	3	0	0	0	0
2	2	1	0	0	0
3	6	0	0	0	0
4	5	1	0	0	0

(e) GLRLM at $\theta = 135^{o}$

Figure 3.8: The four grey-level run-length matrices of a digital image are derived from four directions: 0°, 45°, 90^{0} and 135^{0}.

Table 3.4: List of 11 textural features derived from GLRLM. The definitions of these textural features are described in Appendix A.

Textural feature	Textural feature
Short Run Emphasis (SRE)	High Grey-Level Run Emphasis ($HGRE$)
Long Run Emphasis (LRE)	Short Run Low Grey-Level Emphasis ($SRLGE$)
Grey-Level Nonuniformity (GLN)	Short Run High Grey-Level Emphasis ($SRHGE$)
Run Length Nonuniformity (RLN)	Long Run Low Grey-Level Emphasis ($LRLGE$)
Run Percentage (RP)	Long Run High Grey-Level Emphasis ($LRHGE$)
Low Grey-Level Run Emphasis ($LGRE$)	

3.4.7 Wavelet Transform Based Textural Features

Wavelet transforms have been introduced since early 20th century by Haar (234) and a numerous variations of the wavelet transforms were developed by different groups of scientists, for example Grossman and Morlet (235), Mallat (236), Daubechies (237), and Meyer (238). The wavelet transforms can be divided into two main categories: the continuous wavelet transform (CWT), computing continuous variables, and the discrete wavelet transform (DWT), considering only discrete values. The wavelets are typically used in one-dimensional (1D) for signal decomposition or two-dimensional (2D) transform for image decomposition. In this thesis, the 2D DWT is applied for analysing MR images of childhood brain tumours. The 2D DWT decomposes an image by using father and mother wavelet functions, incorporating with low and high pass filters to analyse image contents at multiscale or multiresolution. The father wavelet is also known as scaling function, $\phi(x)$ and the mother wavelet is also called wavelet function, $\psi(x)$. The scaling function produces a moving average function or a trend of data, whereas the wavelet function measures the differences between the original signal and the average derived from the father wavelet or the fluctuation of data.

The wavelet transform of a signal $f(x)$ is obtained by the convolution of the signal with a scaling function and wavelet function $\psi_{s,t}(x)$ as described in equation (3.43).

$$\langle f(x), \psi_{s,t}(x) \rangle = \int_{-\infty}^{\infty} f(x)\psi_{s,t}(x)dx \tag{3.43}$$

The function $\psi_{s,t}(x)$ is used to create a wavelet family by dilating ψ with a scale factor s and translating ψ by a translation factor t as shown in equation (3.44). The dilation adjusts the size of the wavelet and the translation locates the wavelet transform.

$$\psi_{s,t}(x) = \frac{1}{\sqrt{s}} w\left(\frac{x-t}{s}\right) \tag{3.44}$$

The energy of wavelet defined as $\|\psi_{s,t}(x)\| = \int |\psi_{s,t}(x)|^2 dx$ is usually normalised to have unit energy.

The wavelet transform is required to satisfy an admissibility condition in equation (3.45) so that it can form a complete and numerically stable representation. The sum over entire interval of wavelet is equal to zero satisfying the admissibility condition as shown in equation (3.46).

$$\int \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \tag{3.45}$$

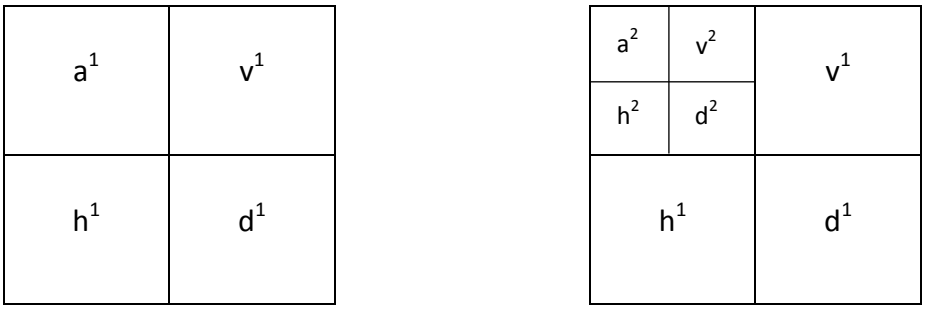
$$\int_{-\infty}^{\infty} \psi(x)dx = 0 \tag{3.46}$$

where $\Psi(\omega) = \int \psi(x)e^{-ix\omega}dx$ is the Fourier transformation of $\psi(x)$.

In 2D DWT, the wavelet decomposition is calculated along the rows and the columns of an image (239). It can be obtained by using the product of two one-dimensional wavelets. First, a wavelet transform is computed on each row to obtain a new image and follows by on each column of a new image. The reverse order (perform wavelet transform on column and then row) can be done and would obtain the same result.

Figure 3.9 graphically demonstrates the image decomposition into 1-scale in Figure 3.9 (a) and 2-scale in Figure 3.9 (b). In the 1-scale image decomposition, d^1 represents diagonal details, h^1 gives horizontal high frequencies, v^1 denotes vertical high frequencies, and the a^1 relates to the lowest frequencies. In the consecutive scale decomposition, the image a^1 are decomposed by using the same low and high filters.

The wavelet transform of an image at the first scale decomposition has the same size as the original image, while at the subsequent decomposition produces next four sub-images with size being twice smaller than the previous scale.



(a) 1-scale decomposition (b) 2-scale decomposition

Figure 3.9: A 2D wavelet decomposition with 1-scale and 2-scale decomposition.

The 2D DWT of three wavelet families: Haar, Daubechies4 (D4) and Symlet4 (sym4), as discussed in the previous section, are considered in this thesis. These wavelet families are orthogonal based of wavelets which offer non-redundant energy map at different scales. Wavelet decomposition at 2 to 4 scales was claimed to provide good texture segmentation (206). The 2D wavelet energies of the aforementioned wavelet families decomposing with 4 scales are used as textural features including energy of approximation coefficients, horizontal detail coefficients, vertical detail coefficients and diagonal detail coefficients. As a result, the Haar, Daubechies 4 and Symlet 4 are

employed and described in the following sections. The comprehensive tutorial of wavelet transform can be found from (240-242).

3.4.7.1 Haar Wavelet

Haar basis, proposed by Haar (234), is one of the earliest and simplest wavelet transform. The Haar transform is a prototype of all other wavelet transforms and is known as Daubechies1 (D1). It decomposes a discrete signal into two sub-signals of half its length (243). One sub-signal is a run of an average or trend of each pair value in a sample by using scaling function, as shown in equation (3.47) and Figure 3.10 (a). Another is a run of difference or fluctuation of each pair value in a sample by using wavelet function, as shown in equation (3.48) and Figure 3.10 (b). The Haar mother wavelet can be derived from the scaling function as $\psi(x) = \phi(2x) - \phi(2x-1)$.

$$\phi(x) = \begin{cases} 1 & 0 \leq x < 1 \\ 0 & otherwise \end{cases} \tag{3.47}$$

$$\psi(x) = \begin{cases} 1 & 0 \leq x \leq 1/2 \\ -1 & 1/2 \leq x < 1 \\ 0 & otherwise \end{cases} \tag{3.48}$$

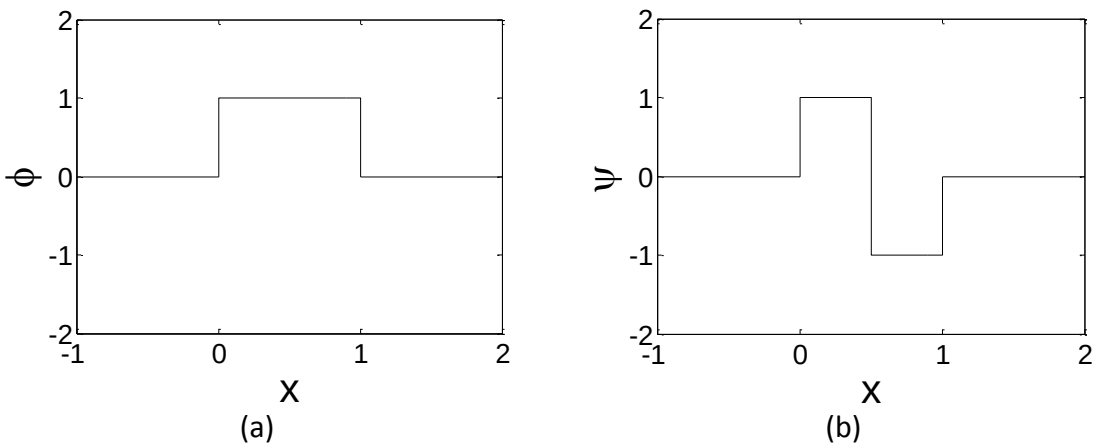


Figure 3.10: Scaling and wavelet functions of Haar Wavelet (a) scaling and (b) wavelet function.

The Haar transform has been useful in denoising diffusion MRI (244). However, it is insufficient to reveal large variation in the high frequency coefficient because pairwise averaging and differentiating processing cannot detect such fluctuation (245). Daubechies wavelet transform was developed to address this issue, as described in the following section.

3.4.7.2 Daubechies Wavelets

Daubechies wavelets adopt the Haar wavelet as a prototype and calculate running averages and differences of signal as Haar transform. However, the Daubechies wavelets provide better detection of high frequency contents than Haar wavelet because of the use of overlapping windows. The DN is the Daubechies wavelet, which has the number of N coefficients. Coefficient N refers to a number of "zero moments" or "vanishing moments" (decay toward low frequencies) equal to half of number of coefficients. The "zero moments" control how the wavelets represent information of a signal or an image in a specific manner. Too high value of N index results in instability. Generally, the even index: $D2$ to $D20$ is used in broad range of applications (237).

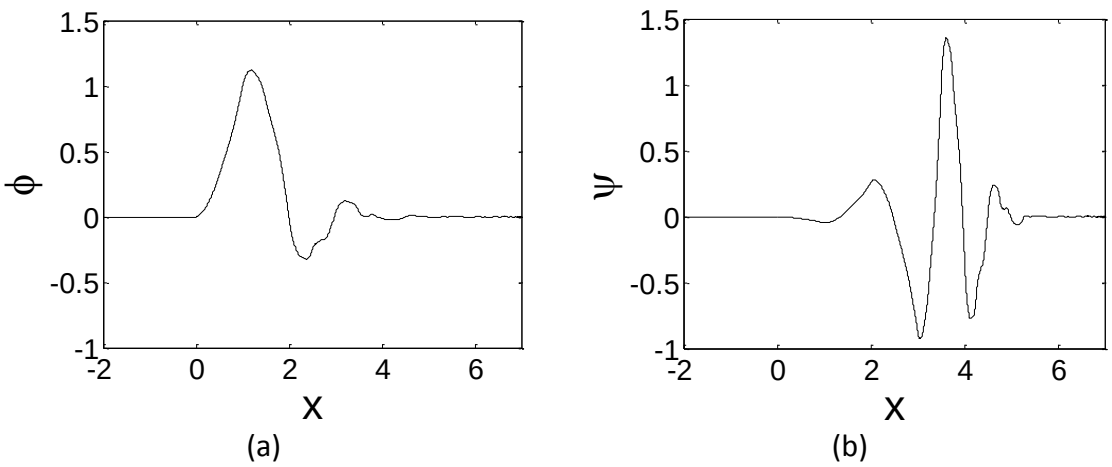


Figure 3.11: Scaling and wavelet functions of $D4$ (a) scaling and (b) wavelet function

Among Daubechies wavelets, $D4$ and $D6$ are popularly applied in medical imaging applications (246). $D4$ with the scaling and wavelet function, depicted in Figure 3.11, is presumed to provide appropriate energy values for MRI based childhood brain tumours.

3.4.7.3 Symlet Wavelets

Symlet stands for Symmetrical wavelet is a modified version of Daubechies wavelet with increased symmetry and is known as Daubechies' least asymmetric wavelets. Apart from the better symmetry of Symlets, the properties of Symlet and Daubechies wavelet families are similar. The Symlet wavelet of order N is denoted as $symN$, where N is the number of vanishing moments for a given support width of $2N$-1 (247). The order of Symlet is related to the quality of image resolution. A higher order degrades the image quality (248) and likely mislead the energy values used as textural features. As aforementioned, wavelet decomposition at 2-4 scales can provide good segmentation result, $sym4$ is therefore adapted in this thesis. A father and mother wavelet of $sym4$ are shown in Figure 3.12 (a) and (b) respectively.

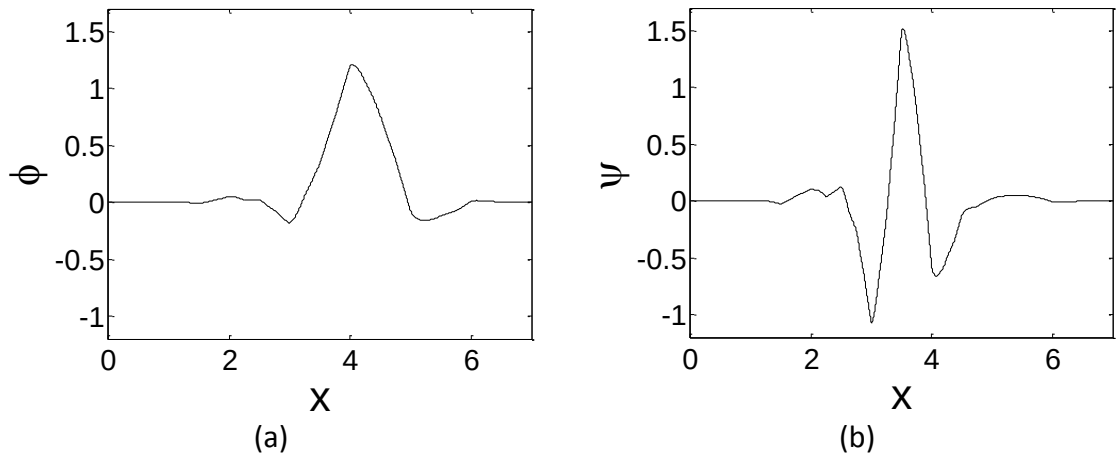


Figure 3.12: Scaling and wavelet functions of $sym4$ (a) scaling and (b) wavelet function.

3.5 Textural Feature Normalisation

Two dimensional MRI produces multiple slices of brain structure. Tumours appear on multiple slices, varying from a few to over twenty slices. The largest visible coverage of tumour is segmented and their textural features are derived for each slice. An average of textural features across slices was used as a representative of each case. These average feature values across all cases vary in different dynamic ranges. Features with large values have more weight on parameters involved in designing the classifier. Therefore, data normalisation is required to scale feature values within predetermined ranges. Each feature was scaled to the same interval range of zero and one, as defined in equation (3.49).

$$x_{i,n,norm} = \frac{x_{i,n} - x_{i,min}}{x_{i,max} - x_{i,min}} \tag{3.49}$$

where $x_{i,n,norm}$ is the normalised value of feature i of case n; $x_{i,n}$ is the value of textural feature i of case n; $x_{i,min}$ is the minimum value of textural feature i of the training data; $x_{i,max}$ is the maximum intensity of textural feature i of the training data; $i \in \{1, 2, 3, \dots, N_f\}$; N_f is the total number of features; $n \in \{1,2,3, \dots, N_{sc}\}$; N_{sc} is the total number of cases.

3.6 Feature Selection

With small numbers of samples and high-dimensional feature space (number of features > 50) (249), the so-called “curse of dimensionality”, resulting in the slow convergence to the solution of any estimator, is often present (250, 251). In addition, irrelevant features

can contribute to unsatisfactory classification accuracy and difficulties in interpreting the feature relations.

Good features should be invariant when there are variations within a class and still reveal important identity among different classes (252). An inappropriate method of feature selection can degrade classifier performance when larger datasets are used (253), therefore an appropriate method of feature selection is important to provide a good feature set.

Feature selection is used to select relevant features to improve the classification accuracy, enhance feature interpretability to better relate the features to the target pathology, as well as, reduce dimensionality to minimise computational load and over-fitting (254). Methods of feature selection can be broadly divided into three main approaches: filter, wrapper and embedded approaches (255). Filter methods do not require learning algorithms but can be unsupervised or supervised, for example PCA and Mutual Information (MI) based feature selection. Wrapper and embedded approaches use a learning algorithm to score feature subsets according to their discriminative selection criteria (256). Wrapper methods can provide more accurate solutions than filter methods (257) because of the optimisation technique of a learning algorithm, however, they generally require higher computation time because the induction algorithm must be evaluated over each feature set. The embedded methods perform feature selection in the process of training. The most common embedded methods are such as Least absolute shrinkage and selection operator (Lasso) and ℓ1-norm regularized Support Vector Machines (255).

PCA is a popular technique, which has been used for MR image based brain tumour classification (19, 258). MI based feature selection has also been widely applied on medical data analysis (259-261). Minimal-Redundancy-Maximum-Relevance (mRMR), proposed by Peng et al. (257), is one of popular MI based feature selection methods, which has been useful for classifying MR image based brain diseases including brain tumours (262-264). In this thesis, the unsupervised PCA method and the wrapper approach based on mRMR and sequential forward selection (SFS) are considered, as explained in the following sections.

3.6.1 Principal Component Analysis

Principal component analysis (PCA) is a feature extraction method which is commonly used for dimensionality reduction. It is an unsupervised feature selection technique which generates a new set of features, called principal components (PCs). Each of the PCs is a linear combination of original variables with different coefficients associated to each original variable. The PCA transforms the original feature set in a way that all the PCs are orthogonal to each other to eliminate redundant information and arrange the PCs in descending order from highest variance (the first PC) to lowest variance (the last PC). The entire set of PCs is the same as the original feature set.

The method of PCA is composed of normalising the original feature set, which has n observations and p variables, to zero mean and one standard deviation shown in equation (3.50).

$$z_{jk} = \left(x_{jk} - m_k\right)/\sigma_k \tag{3.50}$$

where $x_{jk} \in \boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$ is the original features; $z_{jk} \in \boldsymbol{Z} = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1p} \\ z_{21} & z_{22} & \dots & z_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{np} \end{bmatrix}$ is the normalised features; $m_k \in \boldsymbol{M} = \begin{bmatrix} m_1 & m_2 & \dots & m_p \\ m_1 & m_2 & \dots & m_p \\ \vdots & \vdots & \ddots & \vdots \\ m_1 & m_2 & \cdots & m_p \end{bmatrix}_{n \times p}$ is the mean of the original features, m_k is the mean of the kth feature; $\sigma_k \in \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1 & \sigma_2 & \dots & \sigma_p \\ \sigma_1 & \sigma_2 & \dots & \sigma_p \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1 & \sigma_2 & \cdots & \sigma_p \end{bmatrix}_{n \times p}$ is the standard deviation of the original feature, σ_k is the standard deviation of the kth feature; $j \in \{1,2, \dots, n\}$; $k \in \{1,2, \dots, p\}$.

The eigenvector $\boldsymbol{V}$ is calculated from the covariance of $\boldsymbol{Z}$ and is used to calculate the PCs. The PCs are the transformed data ($\boldsymbol{Y}$), computed from the multiplication of $\boldsymbol{Z}$ with the transpose of eigenvectors, as defined in equation (3.51).

$$\boldsymbol{Y} = \boldsymbol{V}^{\boldsymbol{T}} \boldsymbol{Z} \tag{3.51}$$

3.6.2 Minimum Redundancy and Maximum Relevance

Minimal-Redundancy-Maximum-Relevance (mRMR) is a filter and supervised based feature selection method. It employs a mutual information technique to select the most relevant features with the true observation classes and reduce the redundancy among selected features concurrently.

The mutual information of the two discrete random variables, $I(X, Y)$ quantifies an amount of mutual dependence of the two discrete random variables, X and Y. Given $x \in X$ and $y \in Y$, their mutual information of the two variables x and y is defined in

terms of their probability density functions $p(x)$, $p(y)$ and joint probability density function, $p(x, y)$.

The probability density function, $p(x)$ of X is the function $f(x)$ that x take on a value in the interval $[a, b]$, as shown in equation (3.52).

$$p(x) = P(a \leq X \leq b) = \int_a^b f(x)dx \tag{3.52}$$

The joint probability density function $p(x, y)$ of X and Y is the function $f(x, y)$ that satisfies the conditions, $f(x, y) \geq 0 \; \forall x, y$ and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f\,(x, y)dx\,dy = 1$.

$$\begin{aligned} p(x, y) = P(X = x \; and \; Y = y) &= P(Y = y | X = x) \cdot P(X = x) \\ \text{or} \quad &= P(X = x | Y = y) \cdot P(Y = y) \end{aligned} \tag{3.53}$$

The mutual information of two variables x and y is

$$I(X, Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) log \frac{p(x, y)}{p(x)p(y)} \tag{3.54}$$

To determine the maximal relevance, the amount of dependency (D) of a feature x_i related to class c is measured. The mutual information is zero if x_i and c are independent, so knowing x_i gives no information about c. The mutual information is one if x_i and c are identical, so knowing x_i provides value of c. Given a feature set S and a cardinality (a number of elements of the set) $|S|$, max-relevance is calculated as in equation (3.55).

$$\max D(S, c), \; D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c) \tag{3.55}$$

In filter based feature selection, the combination of the most relevant features does not necessarily produce a good classification result because the features can be highly correlated and provide a limited power of discrimination (254). In order to reduce the

dependency of individual features on each other, the value of min-redundancy is calculated from the mutual information of pairwise features (x_i, x_j), as shown in equation (3.56).

$$\min R(S),\ R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i, x_j) \tag{3.56}$$

The combining of two criteria is called minimal-redundancy-maximal-relevance (mRMR). There are two mRMR schemes: mutual information difference (MID) (3.57) and mutual information quotient (MIQ) (3.58).

The MID scheme is the subtraction of redundant score from relevant score.

$$\max(D, R) = D - R \tag{3.57}$$

The MIQ scheme is the division of relevant score over redundant score.

$$\max(D, R) = D/R \tag{3.58}$$

Ding and Peng (254) recommended that for discrete features, MIQ gives a smaller error compared to MID. However, Gulgezen et al. (265) showed a contradictory result that both MID and MIQ provide comparable accuracy but MID gives a more stable feature set. They also suggested the use of the weighting parameter α for different datasets for better stability and accuracy as shown in equation (3.59).

$$\max(D, R) = \alpha D - (1 - \alpha) R \tag{3.59}$$

According to Gulgezen et al. (2009)'s study, the MID scheme of mRMR defined by equation (3.60) is used as the main feature selection criteria in this thesis. The first mth features are selected from the entire feature set, defined as follows.

$$\max(D, R) = \max_{x_i \in X - S_{m-1}} \left[I(x_i, c) - \frac{1}{m-1} \sum_{x_j \in S} I(x_i, x_j) \right] \tag{3.60}$$

3.6.3 Sequential Forward Selection

Sequential Forward Selection starts with no variables in the model and sequentially adds variables to an empty candidate set until the addition of further features does not increase the classification accuracy. If any candidate subset provides the same accuracy, the first combination feature subset containing higher discriminative power is selected in order to reduce over-fitting, as demonstrated in Figure 3.13. Let V_f be a set of feature space, $f_i \in V_f, i \in \{1,2,3..,|V_f|\}$.

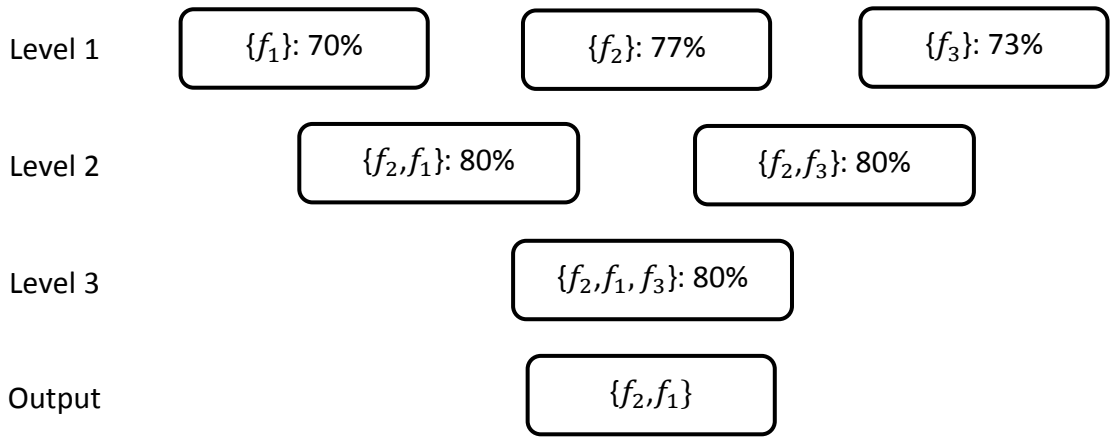


Figure 3.13: A diagram of sequential forward selection. The process starts with one feature variable in the feature set and records the feature giving the highest classification accuracy. Add a new feature into the feature set to find a new value of classification outcome. Repeat the process to the last feature variable. Then the output is the feature set giving the highest classification accuracy and lesser number of features.

3.7 Classification

A number of standard classification methods have been applied to differentiate MRI-based brain tumour types for example, two of the most popular methods: support vector machines (SVMs) and neural networks (19, 30, 227, 258, 266-268). Compared with neural

networks, SVMs were shown to be a better classifier for MRI-based paediatric brain tumour types (184, 266). As a result, a SVM method is selected for single modality MRI based TA.

In order to pursue multimodal MRI based TA, a classifier that can provide probability values, is used instead of the SVM. A multinomial logistic regression is one of popular probability classification algorithms and is a common used technique for medical data analysis (181, 191). As a result, this method is employed as a classifier for multimodal MRI based TA.

3.7.1 Support Vector Machines

SVMs were developed by Cortes and Vapnik (269) and are considered as state-of-the-art supervised learning models used in various applications (270). Basically, the SVMs build the maximum marginal separation line from a set of training samples to separate two classes. By using a kernel function, it maps the original features into higher dimensional space where it computes a hyperplane that maximises the distance from the hyperplane to the training data in each class. The SVMs predict unlabelled test data by mapping it into the feature space to indicate the side of the separating plane where the test data lays. The LIBSVM tool (271) with 'one-against-one' multiclass approach and linear kernel function is adapted. The 'one-against-one' method yields $k(k-1)/2$ binary models, where k is the number of classes. The linear kernel is employed because mapping high dimensional data into another space can linearly separate different classes. It also offers lower computational load compared to other kernels. The hyperplane and decision function can be explained mathematically as follows.

Given n_s training samples, $\{x_i, y_i\}$, where $i = 1, \dots, n_s$, each sample has c features, $\boldsymbol{x_i} \in R^C$ in two classes and a label vector $y \in \{-1, +1\}$. The hyperplane that separates a binary data and the maximum marginal separation lines are defined by (3.61) – (3.63) respectively.

$$\boldsymbol{w} \cdot \boldsymbol{x} + b = 0 \tag{3.61}$$

$$\boldsymbol{w} \cdot \boldsymbol{x_i} + b \geq +1 \quad when\ y_i = +1 \tag{3.62}$$

$$\boldsymbol{w} \cdot \boldsymbol{x_i} + b \leq -1 \quad when\ y_i = -1 \tag{3.63}$$

where $\boldsymbol{w}$ is the weight vector orthogonal to the hyperplane, b is a bias translating the hyperplane away from the origin.

In LIBSVM, the maximum margin classifier is the discriminant function that maximises the geometric margin $\frac{1}{\|\boldsymbol{w}\|}$ which is equivalent to minimising $\| \boldsymbol{w} \|^2$ defined as

$$\underset{w,b}{\text{minimise}} \quad \frac{1}{2}\boldsymbol{w^T w} + C\sum_{i=1}^{Ns} \xi_i \tag{3.64}$$

$$subject\ to: \quad y_i(\boldsymbol{w^T}\phi(\boldsymbol{x_i}) + b) \geq 1 - \xi_i, \ \ \xi_i \geq 0 \tag{3.65}$$

where $\xi_i$is a slack variable that allows an example to be in the margin, $\phi(\boldsymbol{x_i})$ maps $\boldsymbol{x_i}$ into a higher-dimensional space, and $C > 0$ is the regularization parameter.

The decision function is generally defined mathematically as:

$$sign(\boldsymbol{w^T}\phi(\boldsymbol{x}) + b) = sign(\sum_{i=1}^{n_{sv}} y_i \alpha_i K(\boldsymbol{x_i}, \boldsymbol{x}) + b) \tag{3.66}$$

where $K(\boldsymbol{x_i}, \boldsymbol{x}) = \phi(\boldsymbol{x_i})^T \phi(\boldsymbol{x})$ is the kernel function used for transforming data into the linearly separable feature space, α_i is a weighting parameter, $\boldsymbol{x_i}$ is the support vector, $\boldsymbol{x}$ is the input vectors mapped into a higher dimensional space by the kernel function, n_{sv} is the number of support vectors.

3.7.2 Logistic Regression

Logistic regression (272), also called a logit model, is a type of regression analysis used for predicting the outcome of categorical (nominal) dependent variables based on one or more predictor variables. Assumptions of logistic regression allow the classification of brain tumours based on MR textural features to be feasible. A dependent variable does not need linear relationship with predictor variables. There is also no requirement for the predictor variables to be interval, normally distributed, or have equal variance within each group (273). It also offers probabilistic values of predicted outcomes, permitting the manipulation of multimodal MR image analysis as described in chapter 8.

3.7.2.1 Binary Logistic Regression

The logistic regression is a linear function of predictor variable(s) and can be seen as a non-linear transform of linear regression. The logit of a dichotomous outcome and one predictor variable x_i is modelled as equation (3.67).

$$logit\ (Y_i = 1) = ln\left[\frac{P(Y_i = 1|x_i)}{1 - P(Y_i = 1|x_i)}\right] = \alpha + \beta x_i \tag{3.67}$$

where $P(Y_i = 1|x_i) = \frac{1}{1+e^{-(\alpha+\beta x_i)}}$ is the probability of the outcome Y_i=1, α is the intercept of the logit against the predictor variable on Y axis, and β is the regression coefficient. The ratio $\frac{P(Y_i=1|x_i)}{1-P(Y_i=1|x_i)}$ is called an odd ratio, explaining the relationship between a predictor and a dependent variable. The relationship between the probability of the outcome and x_i has a sigmoidal distribution, while $logit\ (Y_i)$ and x_i is linearly related.

The value of regression coefficients directs the relationship between the predictor variable x_i and $logit\ (Y_i)$; the larger the β value, the greater association of x_i with dependent variable Y_i.

3.7.2.2 Multinomial Logistic Regression

When a nominal outcome variable has more than two categories, the model is called multinomial, polychotomous or polytomous logistic regression (274). The multinomial logistic regression (MLR) model explains the relative chance of being in one category versus being in the reference category K. The multinomial logit model of K category outcomes can be described by using a linear combination of j predictor variables as shown in equations (3.68) - (3.70).

$$logit\ (Y_i = 1) = ln\left(\frac{P(Y_i = 1|x_i)}{P(Y_i = K|x_i)}\right) = \alpha_1 + \beta_{11}x_1 + \beta_{12}x_2 + \cdots + \beta_{1j}x_j \tag{3.68}$$

$$\begin{aligned} logit\ (Y_i = nj) &= ln\left(\frac{P(Y_i = nj|x_i)}{P(Y_i = K|x_i)}\right) \\ &= \alpha_{nj} + \beta_{nj,1}x_1 + \beta_{nj,2}x_2 + \cdots + \beta_{nj,j}x_j \end{aligned} \tag{3.69}$$

$$\begin{aligned} logit\ (Y_i = K-1) &= ln\left(\frac{P(Y_i = K-1|x_i)}{P(Y_i = K|x_i)}\right) \\ &= \alpha_{k-1} + \beta_{k-1,1}x_1 + \beta_{k-1,2}x_2 + \cdots + \beta_{k-1,j}x_j \end{aligned} \tag{3.70}$$

where α_{nj} are the intercept of the logit against the predictor variable x_j on Y axis, while keeping other predictor variables constant; $\beta_{nj,j}$ are the regression coefficients; x_j is the jth predictor variables; $nj \in \{1, \ldots, K-1\}$.

Multinomial logit models compute the probability related to each category as defined in equations (3.71) – (3.73).

$$P(Y_i = 1|x_i) = \frac{e^{\beta_1 x_i}}{1 + \sum_{k=1}^{K-1} e^{\beta_k x_i}} \tag{3.71}$$

$$P(Y_i = nj'|x_i) = \frac{e^{\beta_{nj'} x_i}}{1 + \sum_{k=1}^{K-1} e^{\beta_k x_i}} \tag{3.72}$$

$$P(Y_i = K|x_i) = \frac{1}{1 + \sum_{k=1}^{K-1} e^{\beta_k x_i}} \tag{3.73}$$

where $P(Y_i = K|x_i)$ is the probability of $Y_i = K$; x_i is a vector of predictor variables; β_k is a vector of beta coefficients; $nj' \in \{1, \dots, K\}$.

The coefficients in the model explain the effects of the predictor variables on the relative risk or the log odds of being in one category versus the reference category K. Considering coefficient β_{11}, the relative log odds of the dependent variable being category 1 versus in category K increases β_{11} times with a one-unit increase in x_1, given all else equal. In terms of probability estimation, the coefficient β_{11} indicates that the probability of the outcome being in category 1 compared to the probability of being in category K increases exp(β_{11}) times for each unit increase in x_1, given all else is held constant.

The function of multinomial logistic regression is employed from MATLAB toolbox. In MATLAB, the coefficients in the logit model are computed by using the iteratively weighted least squares algorithm (275) as the maximum likelihood method, which is equivalent to iteratively searching for the smallest possible deviance or best fit between the observed and predicted. Once it has found the best solution, it provides a final value for the deviance.

3.7.2.3 Goodness of Fit

To evaluate the performance of the logistic regression model, the "goodness of fit" (276), statistical tests of predictor variables and classification accuracy are examined. The "goodness of fit" of the observed and the predicted values is evaluated using deviance. A larger deviance represents the deviation of observed values from the predicted values and implies poorer fit to the model and vice versa. Once a suitable functional form of a model is derived using feature selection and cross-validation techniques, deviance can be employed to measure how well the model fits observed values. Deviance values can also be used to measure the relative level of fit (to observed values) of models derived from different sets of textual features.

The statistical tests of predictor variables are computed using the Wald chi-square statistic which is calculated as $\left(\frac{B}{S.E.}\right)^2$, where B is a beta coefficient and $S.E.$ is a standard error. The classification accuracy of each class and overall accuracy is also calculated and presented in the form of the confusion matrix, tabulating a number of correct predicted cases and incorrect predicted cases against the true observed cases.

3.8 Validation for Classification

Cross-validation and bootstrapping are the two main approaches in estimating classifier performance. K-fold cross-validation partitions data into k segments or folds with equal or nearly equal segment size. The k iterations are performed while each fold is held-out as a test set and k-1 folds are used for learning. In genetic data analysis, 10-fold and 5-fold cross validation are two recommended techniques (277, 278), however, this approach

requires a large sample size to provide a reliable prediction error rate. Leave-one-out cross-validation (LOOCV) is a special case of k-fold cross-validation where k is equal to the number of instances in the data. The LOOCV approach is used when the smallest group size is at least five times the number of predictor variables (279). Amongst classifier validation methods, LOOCV method has a propensity to produce the highest prediction accuracy for the classification problems due to its capability to process almost all of the available data for training the classifier (280). It was also demonstrated that LOOCV has high variance (281). The bootstrapping methods were introduced to reduce bias raised from LOOCV.

A bootstrap estimates prediction error by re-sampling the training data, x_t. This approach randomly chooses n elements of x_t with replacement and repeats the process B times to obtain B bootstrapping samples. Due to the replacement process, some elements of x_t can be selected multiple times and some element may not be selected at all. In this thesis, 50 bootstrap replications were used as recommended by Efron and Tibshirani (282). The .632 estimator is a classical bootstrap method, however, this technique suffers from downward bias (283). The .632+ estimator is an improved version and aims to correct the upwardly biased leave-one-out bootstrap (LOOB) and downwardly biased .632 estimator (282).

The validation methods based on LOOCV, LOOB, .632 estimator and .632+ estimator are considered, in this thesis. These validation techniques can be mathematically explained as follows:

Given the observation data $x_t = \{x_1, x_2, x_3, \dots x_n\}$, the classifier rule C, the discrepancy between the predicted value $C(x_t)$ and response y.

$$F(y, C(x_t)) = \begin{cases} 0 & if\ C(x_t) = y \\ 1 & if\ C(x_t) \neq y \end{cases} \tag{3.74}$$

The LOOCV error, $\widehat{Err}^{(cv1)}$ is defined in equation (3.75).

$$\widehat{Err}^{(cv1)} = \frac{1}{n}\sum_{i=1}^{n} F(y_i, C(x_{t(i)})) \tag{3.75}$$

where $x_{t(i)}$ is the training set with the ith case being removed.

The LOOB error, $\widehat{Err}^{(1)}$ is shown in equation (3.76).

$$\widehat{Err}^{(1)} = \frac{1}{B}\left(\sum_{b=1}^{B} \frac{1}{n}\sum_{i=1}^{n} F(y_i, r_{x^b}(t) C(x_{t(i)}))\right) \tag{3.76}$$

The .632 estimator is defined in equation (3.77).

$$\widehat{Err}^{(.632)} = .368 \cdot \overline{err} + .632 \cdot \widehat{Err}^{(1)} \tag{3.77}$$

The .632+ estimator is defined as in equation (3.78).

$$\widehat{Err}^{(.632+)} = (1 - \widehat{\omega}) \cdot \overline{err} + \widehat{\omega} \cdot \widehat{Err}^{(1)} \tag{3.78}$$

where $\overline{err}$ is the apparent error and $\widehat{\omega}$ is the weighting parameter defined as follows:

$$\overline{err} = \frac{1}{n}\sum_{i=1}^{n} F(y_i, C(x_t)) \tag{3.79}$$

$$\widehat{\omega} = \frac{.632}{1 - .368\hat{R}} \tag{3.80}$$

The .632+ estimator requires the relative over-fitting rate, $\hat{R}$ (equation (3.81)) and the no-information error rate, $\hat{\gamma}$ (equation (3.82)).

$$\hat{R} = \frac{\widehat{Err}^{(1)} - \overline{err}}{\hat{\gamma} - \overline{err}} \tag{3.81}$$

The no-information error rate is estimated based on assumption of independence between the class labels and the feature vectors of the training set. The no-information error for multicategory generalization is defined as:

$$\hat{\gamma} = \sum_{i=1}^{n}\sum_{j=1}^{n} F(y_i, C(x_t))/n^2 \tag{3.82}$$

3.9 Over-Fitting in Pattern Recognition

In pattern recognition, a statistical model is presented with a training dataset with the aim of obtaining useful descriptive features that apply to the larger population from which the training sample was drawn. The over-fitting problem arises when the model incorporates random artefacts or noise from the training sample that does not generalise to the overall population (284). Such a model would perform poorly on a larger test dataset.

Over-fitting is likely to occur where the number of parameters in the model is large relative to the size of the training set (285). In this study, given the small sample sizes, we have used two approaches to reduce the risk of over-fitting. First we employ dimensionality reduction techniques in feature selection to build parsimonious models. Second, we employ cross-validation and bootstrapping techniques which use partitioning to generate multiple training sets. In terms of estimation of classification error rate, bootstrapping and cross-validation statistics evaluate the stability of the classifier over different datasets thereby providing a more accurate reflection of its performance on a larger dataset.

3.10 Statistical Measures

Apart from classification accuracy, performance of the processing methods used in this thesis is measured by the common techniques named sensitivity, specificity, and hypothesis test for two proportions.

3.10.1 Sensitivity and Specificity

The sensitivity and specificity of binary classification test is defined in the same way as the sensitivity and specificity used in evaluation of brain tumour segmentation, in section 3.3.3.2. In brain tumour segmentation image pixels are considered as samples. However, in binary classification, the study cases are used as samples. The sensitivity indicates the positive cases that the method correctly predicted, whereas the specificity indicates the negative cases that the method correctly predicted.

3.10.2 Hypothesis Test for Two Proportions

In order to compare classification accuracy based on significance level, a comparative statistical hypothesis testing is employed. A commonly used method is the comparison of proportion test by estimating the correctly predicted cases samples p_0 and p_1 from sample size n_0 and n_1 respectively. The statistical significance of a difference between two proportions is calculated by equation (3.83).

$$z = \frac{|p_1 - p_0|}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_0}\right)}} \tag{3.83}$$

where $\bar{p} = \frac{x_1+x_0}{n_1+n_0}$ with x_0 and x_1 representing the number of correctly allocated cases in the classifications of samples of size n_0 and n_1 respectively (286, 287).

In this thesis, we use two-tailed Z-test and consider a result to be significant at two confidence intervals. One is at 95% confidence interval (5% significance level); if a p-value is less than 0.05, the test is considered significance. Another is at 90% confidence interval (10% significance level); if a p-value is less than 0.1, the result is considered significance.

3.11 Comparative Studies

In this section, we discuss the methods used in the two recent published papers from Vlachou et al. (19) and Rodriguez Gutierrez et al. (30) in comparison to the methods used in our work.

Vlachou et al. (19) classify the posterior fossa tumours based on TA of pre-contrast T1, T2 and combined textural features from pre-contrast T1 and T2 images, using the MaZda tool. They show that individual pre-contrast T1 and T2 based TA provide higher sensitivity and specificity than the combined textural features obtained from these images. Three main differences are observed between our and their studies.

The first difference relates to the region of interest used. Vlachou et al. (19) use manual segmentation in order to exclude large cystic regions and attempt to use mostly solid tumours. We intend to develop a semi-automatic technique for segmenting tumour regions with the largest possible abnormal region, comprising for example solid and cystic regions.

The second difference is the textural features used. Vlachou et al. (19) use the MaZda tool to generate textural features based histogram, gradient map, GLCM at d = 5, GLRLM, Haar and autoregressive models. The number of grey levels used was not stated in their study. Our study use all the above texture analysis methods except autoregressive models and additionally we use wavelet transform based Dau4 and Sym4. The GLCM and GLRLM are calculated using nine grey levels. The Dau4 and Sym4 based textural features provide complementary information to the Haar based textural features as reviewed in section 3.4.

The third difference between the studies is the use of classification methods. Vlachou et al. (19) compare the performance of classification based on probabilistic neural network (PNN) and linear discriminant analysis (LDA) and show that probabilistic neural network produces higher sensitivity and specificity. In this thesis, we intend to use a SVM and multinomial logistic regression for single modality and multimodal MRI based TA respectively. The discriminative criteria between the two techniques are different, resulting in different classification outcomes. A SVM is chosen instead of a neural network based on the results of similar studies (184, 266).

The second publication we discuss is the work of Rodriguez Gutierrez et al. (30) where they use ADC-based histogram analysis to differentiate single-centre dataset of paediatric posterior fossa tumours. The 25th and 75th percentile and skewness of ADC based histogram analysis are reported to achieve high classification accuracy and outperform post-contrast T1 and T2 based features classification. Three main techniques: intensity

normalisation, texture analysis and defining region of interest differ between the two studies.

For intensity normalisation, they scale data intensity to the mean value of normal-appearing white matter from two small ROIs drawn bilaterally above the ventricles, whereas we intend to choose experimentally between decile-based pricewise linear transform and linear normalisation. Our methods are better suited for automating intensity normalisation.

For texture analysis, they use a smaller set of manually selected features which are input directly into the classifier. In our study, we extract a large set of initial features which are input into a feature selection algorithm which automatically selects a smaller, relevant set of features for classification. Starting with a richer set of descriptive features increases the likelihood of ultimately using a more relevant set of features in classification and thereby the classification accuracy.

For defining the region of interest, they define tumour regions based on post-contrast T1 registered onto T2 to include low contrast tumour or necrotic tissue, but exclude peripheral blood vessels adjacent to enhancing tumour and exclude perilesional oedema. Our tumour regions include a larger abnormal region comprising for example solid tumour and cyst, observed on T2. This is also a result of our preference for using a semi-automatic segmentation technique.

3.12 Summary

The chapter presented a comprehensive literature review in order to select appropriate techniques and parameters for the automated processing pipeline. The processing framework involves pre-processing analysis, tumour segmentation, texture analysis, texture normalisation, feature selection and classification. The theoretical background of chosen methods including evaluation methods for each step was described.

CHAPTER 4 METHODOLOGY

This chapter presents the methodology of the overall processing methods, supporting the three objectives of section 1.2. It describes the four experiments and the quality control of datasets used in this thesis.

4.1 Introduction

Three main objectives were defined to achieve the aim of this thesis articulated in section 1.2. The approach taken to achieve those objectives resulted in the integration of several methods to enhance diagnostic classification performance of paediatric brain tumours based on textural features extracted from conventional and diffusion MR images. Four main experiments were designed to evaluate the methods and optimise the processing outcome. This chapter explains the motivation, presents flowcharts of the experiments and describes methods of the entire processing pipeline. The overview of the entire workflow of this processing framework is shown in Figure 4.1. The four experiments: pre-processing analysis for brain imaging data, brain tumour segmentation, classification of brain tumours based on TA of single modality MRI and classification of brain tumours based on TA of multimodal MRI are explained in section 4.2 to 4.5. The dataset and quality control of datasets used in this thesis is described in section 4.6 and 4.7. The literature review and theoretical background of the related subjects are described in chapter 3.

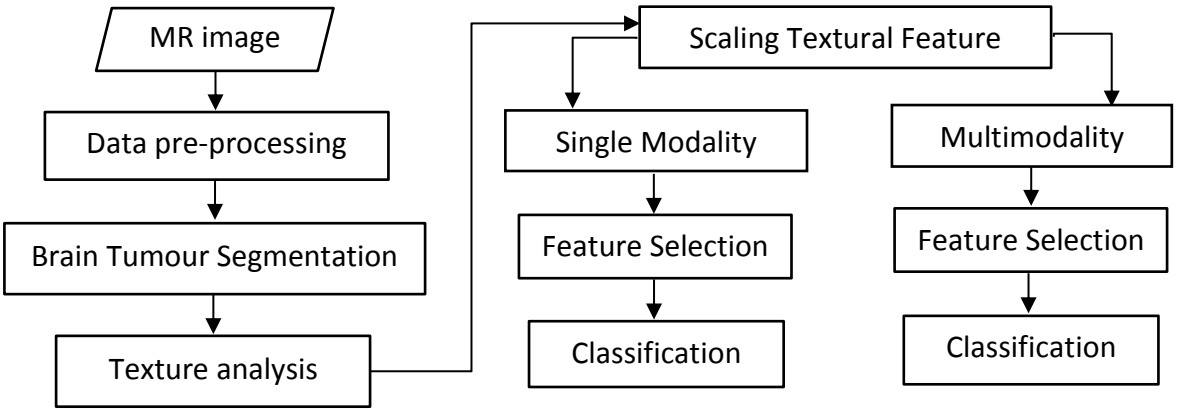


Figure 4.1: Processing pipeline of texture analysis of childhood brain tumour based on single modality and multimodal MRI.

4.2 Experiment 1: Pre-Processing Analysis for Brain Imaging Data

Pre-processing analysis is an important procedure in order to improve the SNR of MR images and therefore enhances TA-based classification accuracy. This pre-processing pipeline also enables multimodal MRI comparison by means of aligning different types of images onto the same spatial location for segmentation as well as standardising intensity ranges. The aim of experiment 1 is to identify the optimal pre-processing methods and associated parameter settings. The work-flow of this experiment is shown in Figure 4.2.

Pre-processing analysis for brain imaging data based MRI has been well established by the FMRIB analysis group, providing a variety of standard neuroimage processing methods. As a result, the processes are executed by running shell scripts on FSL tool, as shown in Appendix B. The theoretical background of individual processing steps is presented in

section 3.2, whereas this experiment details are presented chapter 5. The five processing methods are described below.

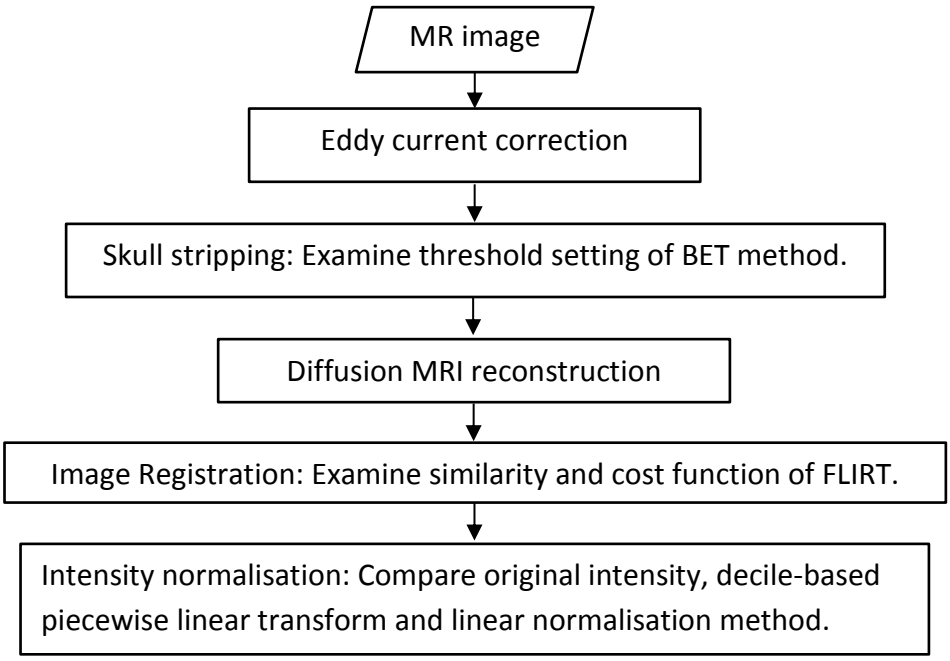


Figure 4.2: Pre-processing analysis of brain imaging data based MRI.

4.2.1 Artefact Removal

As aforementioned in section 3.2.1, the B_0 field mapping and eddy current correction are recommended artefact removal for diffusion MRI. However B_0 field mapping is not applied in this work because the fieldmaps are not available in the CCLG database. To compensate for this, image registration is conducted to reduce motion artefacts and field inhomogeneity. The eddy current correction is applied for diffusion MRI dataset by using FMRIB's Diffusion Toolbox (288). This tool corrects eddy current distortion by using affine registration to a reference volume.

4.2.2 Skull Stripping

The skull stripping based on BET method is used. The choice of this method is explained in section 3.2.2. Due to the diversity of resolution and contrast, suitable threshold values are selected for each MR-based image type. Threshold value is initially set to 0.3. Final choice of threshold value can be changed depending on skull stripping result.

4.2.3 Diffusion Magnetic Resonance Image Reconstruction

ADC maps are reconstructed from either DWI or DTI, subjecting to data availability from CCLG database. ADC map is computed by using equation (2.25) in section 2.2.2. Most of DWI datasets consisted of two b-values: b_0 and b_1 of 0 and 1000 s/mm^2 respectively. In this work, however, three b-values are found in some cases, so the average of the ADC values is calculated, i.e. the first ADC map is obtained from b_1 and b_0; the second ADC map is obtained from b_2 and b_0; the average ADC values are the average of the first and second ADC maps.

FA and MD maps are calculated from DTI dataset as shown in section 2.2.3. The DTI dataset from CCLG database are acquired with the number of scanning directions, ranging from 7 to 42 directions.

4.2.4 Image Registration

In order to align intermodal brain imaging data within the same subjects, linear registration with 6 DOF, correcting rotation and translation of the head during scanning, is recommended by Mark Jenkinson, head of structural image analysis and physics

modelling at FMRIB centre (Jenkinson M., personal communication, July 10, 2012). FMRIB's Linear Image Registration Tool - FLIRT (95, 120), a popular tool used in many studies (200, 289), is employed in this thesis. Similarity and cost function is an important factor for mapping two images, as explained in section 3.2.3. In order to select a suitable mapping function, the correlation ratio (123), a default similarity function of FLIRT, is compared against a cost function named normalised mutual information.

As reviewed in section 3.4, T2 provide the most sensitive imaging to pathology. Consequently, T2-weighted images are selected as reference images for registering multimodal MRI within the same subject.

4.2.5 Intensity Standardisation

Pixel intensity values of brain tissue are standardised to the same intensity interval based on linear normalisation and decile-based piecewise linear transform method. In linear normalisation, brain tissue of MRI and diffusion MRI are normalised to the interval of the standard scales [0,4095] and [0,1] respectively. The intensity at 2% and 98% percentiles of the overall image intensity range for each study case is used as the minimum and maximum robust intensity. For training stage in decile-based piecewise linear transform, the tail of the histogram is avoided and so the minimum and maximum percentile value of the overall intensity range is obtained at 2% and 98%. An effective performance of the methods is compared by calculating the sum of mean difference square, described in section 3.2.4.

4.3 Experiment 2: Brain Tumour Segmentation

Segmentation refers to the separation of background from the region of interest in an image and is an important step in differentiating brain tumour types. Experiment 2 has two stages. In the first stage, we identify which MR-based image type can better partition tumour from non-tumour region based on a proposed segmentation method. In the second stage, using the chosen image type from the first stage, we evaluate the semi-automatic segmentation method against manual segmentation. A critical review of segmentation techniques, presented in section 3.3, is needed to understand relative merits of different techniques in order to select the most suitable existing technique or to postulate a novel technique for brain tumour segmentation. The evaluation methods are reviewed and discussed in section 3.3.3, whereas the details of experiment 2 are described in chapter 6.

4.3.1 Combined MNcut and GVF Snake Hybrid Method

The proposed algorithm is a hybrid semi-automatic segmentation technique, combining the MNcut and the GVF snake method (Figure 4.3). This method is hypothesised to improve the performance of the GVF snake method by reducing the number of input parameters and the likelihood of the snake progressing toward the wrong edges with the initial boundary produced by the MNcut technique. MR images displaying abnormal tissues are cropped to focus on the tumour region and to reduce the computational load for MNcut and GVF snake processing. Initially, tumour images are segmented by using the MNcut method. Sub-partitions are manually selected to form a complete tumour

boundary. This boundary constitutes an initial contour for the GVF snake method to capture the target edges.

When using the MNcut method, the K value (a number of segments) is defined by the user to obtain a segmentation result of acceptable quality. For the GVF snake method, appropriate parameter values of the deformable model were selected partially based on He et al. (290) and experimental results on our dataset.

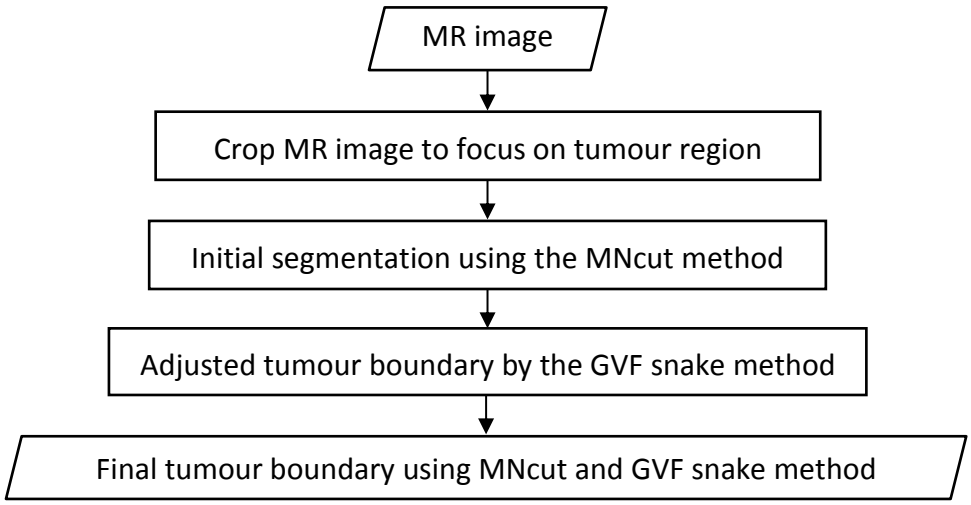


Figure 4.3: The proposed segmentation algorithm based on the combined MNcut and GVF snake method.

4.3.2 Evaluation of Segmentation

Four types of structural MRI are considered; T2, FLAIR, pre-contrast T1, post-contrast T1. In order to compare the segmentation results of four different types of MR images, each image type has to be acquired at the same spatial location (the same slice position and angulation). However, the resolution and especially the slice spacing of each MR sequence is sometimes acquired with different parameter settings. In experiment 2, the most similar plane of each MR sequence was selected and, as such, all images are

assumed as comparable. The original sequences are used instead of co-registered images, because the performance of the segmentation method is required to be examined on the original spatial resolution of each MRI sequence.

4.3.2.1 Five-Point Likert Scale of Evaluation

Three slices displaying a brain tumour region (corresponding to approximately the tumour's middle axial plane) are used for evaluation based on the five-point Likert scale of evaluation criteria (291), which is a commonly used rating scheme. In this study, each item of the scale has a meaning associated to it, as tabulated in Table 4.1. The samples based on these scales were rated independently by the four evaluators.

Table 4.1: Meaning of scores for the five-point Likert scale evaluation form.

Score	Meaning
Strongly agree 5	The segmentation is done very satisfactorily. No change is required (81%-100%).
Agree 4	The segmentation is done satisfactorily. Little change is required (61%-80%).
Neutral 3	The segmentation is done moderately well. Moderate change is required (41%-60%).
Disagree 2	The segmentation is done poorly. Large change is required (21%-40%).
Strongly disagree 1	The segmentation is done very poorly. Exhaustive change is required (0%-20%).

In this thesis, the two-way ANOVA with mixed average measures of samples, denoted as ICC(3,k) is adapted. This form of ICC assumes that each subject was assessed by the same group of evaluators, but these particular evaluators are the only evaluators of interest and evaluators are a fixed effect on this evaluation study. Moreover, the average of k

rating and the mean of ratings are taken into account. This ICC function is implemented by Kevin Brownhill and available at File Exchange – MATLAB Central. In addition to the ICC analysis, the mean and standard deviation of the rating scores are also reported.

4.3.2.2 Supervised Evaluation

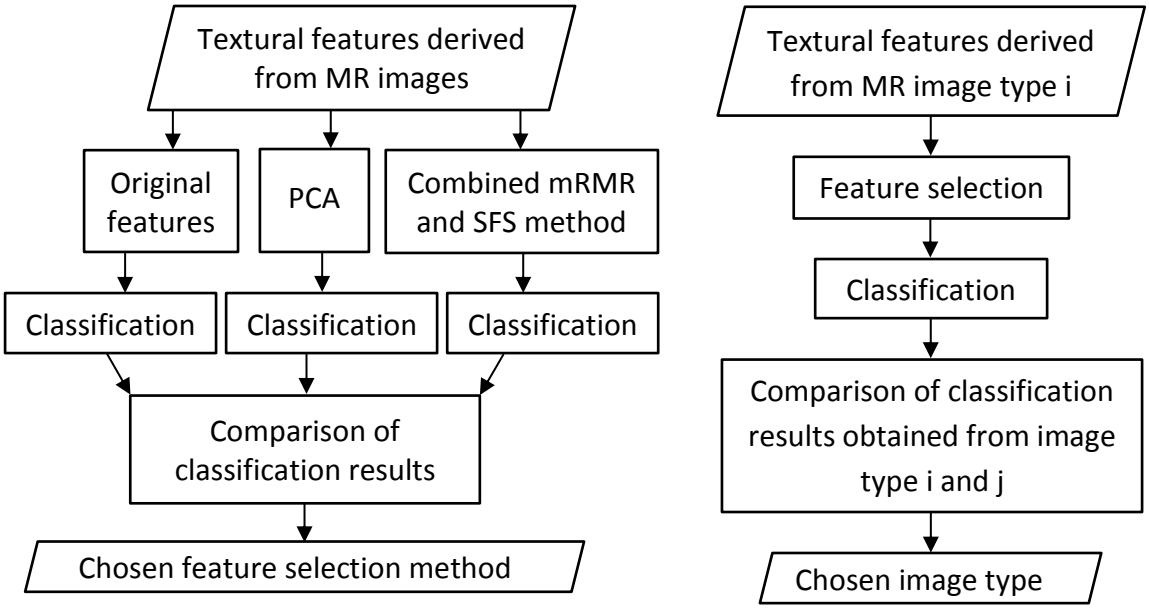
The MRI sequence gaining the highest average score from the five-point Likert scale evaluation is further evaluated based on the supervised evaluation approach. The manual segmentation is conducted by interpolating selected adjacent points on the tumour boundary. The proposed segmentation method was evaluated against the manual segmentation by measuring the sensitivity, specificity, modified Hausdorff distance, and the volume overlap ratios including dice coefficient as described in section 3.3.3.

4.4 Experiment 3: Classification of Brain Tumours Based on Texture Analysis of Single Modality MRI

Experiment 3 has two objectives. The first objective is identifying a feature selection method that performs well for this dataset. Using the chosen feature selection method, the second objective is ranking the classification results of individual MR image types. The workflows of the two steps are shown in Figure 4.4. The theoretical background of pattern recognition methods and validation techniques are reviewed in section 3.4 to 3.10. The details of experiment 3 are described in chapter 7.

In the first part, the two feature selection methods: PCA and the combined mRMR and SFS method are compared using the following methodology (Figure 4.4(a)). First raw

textural features are extracted from seven types of MR-based images: T2, FLAIR, pre-contrast T1, post-contrast T1, ADC, MD and FA. For each image type a set of features are selected using the PCA and the combined mRMR and SFS method. Those chosen features are then input into the same classifier (SVM with a linear kernel). The feature selection method that performed better (in terms of classification accuracy) across all MR image types is deemed to be the more suitable technique. The performance evaluation employs a hypothesis test for two proportions.



(a) (b)

Figure 4.4: Classification of brain tumours based on TA of Single Modality MRI. The workflows of (a) identifying a feature selection method and (b) ranking MR-based image types based on classification accuracy.

The second objective is ranking the MR-based image types by the classification accuracy (Figure 4.4(b)), so that the top-ranked image types may be used in the generation of a multimodal feature set to be used in experiment 4. The methodology for experiment 3 involves computing a classification accuracy figure for each image type (across the same number test cases) keeping the feature selection and classification methods constant. The

classification accuracy for each image type is estimated using well known cross-validation and bootstrapping techniques as well comparative statistical testing. Specifically LOOCV, LOOB, .632 estimator, .632+ estimator methods are used.

4.4.1 Principal Component Analysis

To examine the classification accuracy of single modality MR image-based TA, the number of PCs against the prediction rate is calculated. The number of PCs that provides the maximum balanced accuracy (an average accuracy obtained from medulloblastoma, pilocytic astrocytoma and ependymoma) is used to report the classification accuracy, sensitivity and specificity of a dataset. The built-in PCA function in MATLAB is used in this thesis. This classification result is also compared with the classification based on the original features and the features selected by a combined mRMR and SFS method.

4.4.2 Combined mRMR and SFS method

As reviewed in section 3.6.2, the MID scheme of mRMR defined by equation (3.60) is employed with the first ten features being selected from the entire feature set. In order to optimise the classification result, the SFS approach is adapted to search the optimal combination within the ten top-ranked features. For scoring a feature in SFS, the accuracy is calculated by averaging the three prediction accuracy values of medulloblastoma, pilocytic astrocytoma and ependymoma. When there are multiple feature sets having the same average accuracy, a feature set with a smaller cardinality is chosen in order to reduce the risk of over-fitting.

4.5 Experiment 4: Classification of Brain Tumours Based on Texture Analysis of Multimodal MRI

Experiment 4 aims to examine whether the TA of multimodal MRI can improve the classification of childhood brain tumours compared to TA of single modality MR techniques, as illustrated in Figure 4.5. The use of MR image types in multimodal MRI is based on the literature review and the results obtained from experiment 3, while the choice of feature selection is derived from experiment 3. Multinomial logistic regression and LOOCV techniques are used as the classifier and error estimator. Therefore the specific methods of experiment 4 are further detailed in chapter 8.

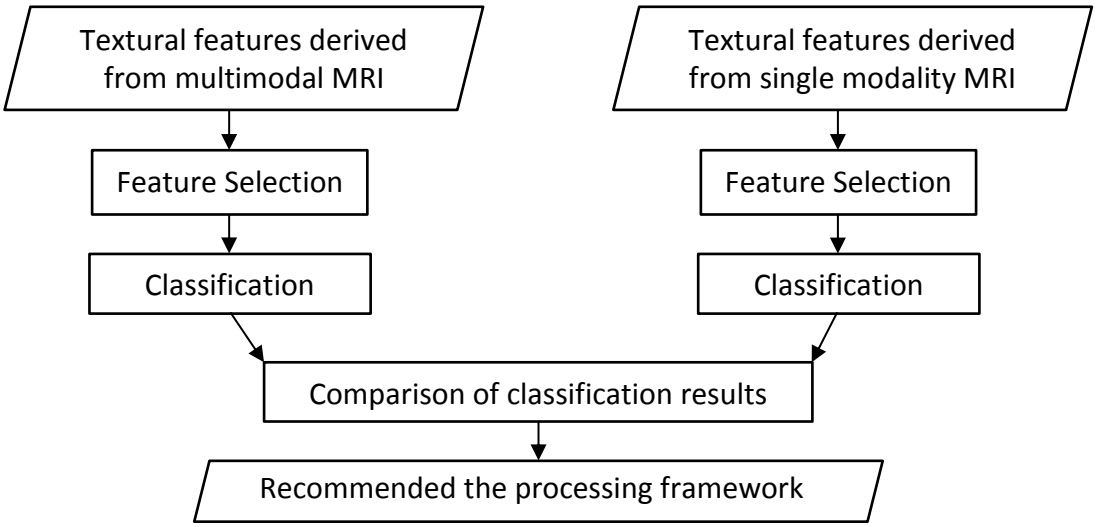


Figure 4.5: Classification of brain tumours based on TA of Multimodal MRI.

4.6 Dataset

In order to maximise the available number of brain tumour study cases, the experiment datasets are acquired from multiple centres, stored in the CCLG database. The use of this

data is governed by a protocol defined by the CCLG Functional Imaging Group. The experiment datasets are anonymous to the author; as a result, there is no requirement for ethics approval of this research. The number of study cases used in each experiment is the maximum number of cases available at the time of the study. The axial multimodal MRI scans including T2, FLAIR, pre-contrast T1, post-contrast T1, DWI and DTI are used in this thesis.

The pre-treatment dataset of embryonal, astrocytic and ependymal tumours are considered. The embryonal tumour type includes classic medulloblastoma (grade IV) and desmoplastic/nodular medulloblastoma (grade IV). Astrocytic tumour type includes pilocytic astrocytoma (grade I). Ependymal tumour type includes ependymoma (grade II) and anaplastic ependymoma (grade III). In this thesis, embryonal and ependymal tumour types are denoted as medulloblastoma (MB) and ependymoma (EP), respectively. Pilocytic astrocytoma is only subtype used from astrocytic tumour type and denoted as PA. These tumours are located in either infratentorial or supratentorial region.

The nomenclature for labelling individual cases is the abbreviation of tumour type followed by trial number, e.g. MB718, PA622 and EP611. The imaging parameters of all study cases and all image types are separated into four experiments, listed in Appendix C.

4.7 Quality Control

Due to sources of variation in the data acquisition, the selection of the dataset for this study is qualified using the following criteria.

1) The dataset with strong effect of ghost or motion artefact were excluded from the study.
2) In order to reconstruct DWI and DTI, b-matrix and b-value files are required by converting the DICOM to NIFTY file format. However, some cases are possibly corrupted and contain no b-matrix and b-value files. These cases are not included into the study.
3) Due to protocol variations at the different centres and multiple slices acquired for the analysis, an image type would be excluded according to the slice spacing criteria. If the registered images have a slice spacing that is different from the T2 image (the reference image for registration and segmentation) by more than 25%, that image type is excluded. The slice difference is calculated from 100 x ($s1$ - $s2$)/($s1$ + $s2$), where $s1$ is a T2 image's slice spacing and $s2$ is the slice spacing of the image being considered.
4) If the skull stripping method fails to perform satisfactorily on any dataset, the original dataset is used for the analysis.

4.8 Summary

This chapter described the methodology of the four experiments employed to answer the research question of section 1.2, whether the multimodal MRI based TA can improve the classification of childhood brain tumours compared to single modality MR technique based TA. The four experiments include pre-processing analysis of brain imaging data, brain tumour segmentation, classification of brain tumours based on texture analysis of

single modality MRI and classification of brain tumours based on texture analysis of multimodal MRI.

CHAPTER 5 EXPERIMENT 1: PRE-PROCESSING ANALYSIS FOR BRAIN IMAGING DATA

This chapter presents the first experiment of the pre-processing framework described in section 4.2, supporting all three objectives of section 1.2. The background theory and literature review of pre-processing pipeline are described in section 3.2. In this chapter, neuroimage processing including examining different techniques and parameter settings are presented.

5.1 Introduction

In experiment 1, MR-based image processing includes artefact removal, skull stripping, diffusion MRI reconstruction, image registration and intensity normalisation. These techniques help to enhance the performance of segmentation, texture analysis and classification that are studied in more detail in chapters 6, 7 and 8. In the following sections we examine the parameter settings and compare the performance of several pre-processing techniques.

As reviewed in section 3.2, a number of software packages are available for analysing brain imaging data based on MR images. FSL is one of the most established and specialised software libraries for analysing brain imaging data based on functional, conventional and diffusion MRI. FSL has robust algorithms, provides compatibility on various operating systems and offers a user-friendly Graphical User Interface. It also provides comprehensive procedures for the pre-processing pipeline including artefacts

removal, skull stripping, registration and DTI map reconstruction. Due to several benefits, FSL is employed for this processing analysis.

5.2 Material

A pre-treatment dataset of paediatric brain tumours based on seven MR-based image types depicting medulloblastoma, pilocytic astrocytoma and ependymoma cases are used in experiment 1. The number of MB, PA and EP cases available at the time this study was undertaken are listed in Table 5.1. The average and range of TR, TE, TI, resolution and slice thickness of the brain imaging data are shown in Table 5.2. The study cases with value of TR, TE, TI, resolution, FOV, slice thickness and magnetic field strength as well as the scanners used in experiment 1 are listed in section C.1, Appendix C.

Table 5.1: Number of MRI and diffusion MRI dataset of three tumour types.

Image type	Medulloblastoma	Pilocytic Astrocytoma	Ependymoma
ADC	25	33	18
MD	19	16	8
FA	19	16	8
T2	25	33	18
Pre-contrast T1	21	24	15
Post-contrast T1	21	29	17
FLAIR	20	19	9

Table 5.2: Imaging parameters for experiment 1. Average and range of TR, TE, resolution and slice thickness are presented in the format of average (min, max).

Image type	TR (ms)	TE (ms)	Resolution (mm)	Slice thickness (mm)
ADC	4961 (2486,8700)	86 (48,161)	[1.4,1.4,5.1] ([0.8,0.8,2.5],[2.4,2.4,8])	4.3 (2.5,5)
DTI	6057 (2486,12000)	78 (48,110)	[1.8,1,8,3.7] ([0.9,0.9,2.5],[2.4,2.4,6.5])	3.4 (2.5,5)
T2	4766 (2660,9140)	87 (14,119)	[0.5,0.5,5] ([0.4,0.4,3.9],[0.9,0.9,8.5])	4.0 (3,5)
FLAIR	9930 (8002,1100)	116 (89,140)	[0.6,0.6,5.2] ([0.4,0.4,3],[1,1,6.6])	4.4 (3,6)
Pre-T1	602 (360,1400)	13 (8.4,23)	[0.6,0.6,5.3] ([0.4,0.4,4.4],[0.9,0.9,7.5])	4.4 (4,5)
Post-T1	557 (8.2,828)	15 (2.3,23)	[0.6,0.6,5.4] ([0.4,0.4,1.9],[0.9,0.9,7.9])	4.2 (0.8,6)

TI of FLAIR = 2499 (2000, 2800) ms.

5.3 Results and Discussion

5.3.1 Artefact Removal

Eddy current correction was applied on DWI and DTI to remove eddy current artefact, induced during acquisition of diffusion MRI. Figure 5.1 shows original, eddy current corrected and the difference of intensity between the two images. Even though the effect of eddy current correction is not apparent from visual inspection, the method is widely adopted to correct for this artefact, which causes geometric distortions. Hence we adopt the eddy current correction method offered by FSL as a verified standard pre-processing step in our processing pipeline.

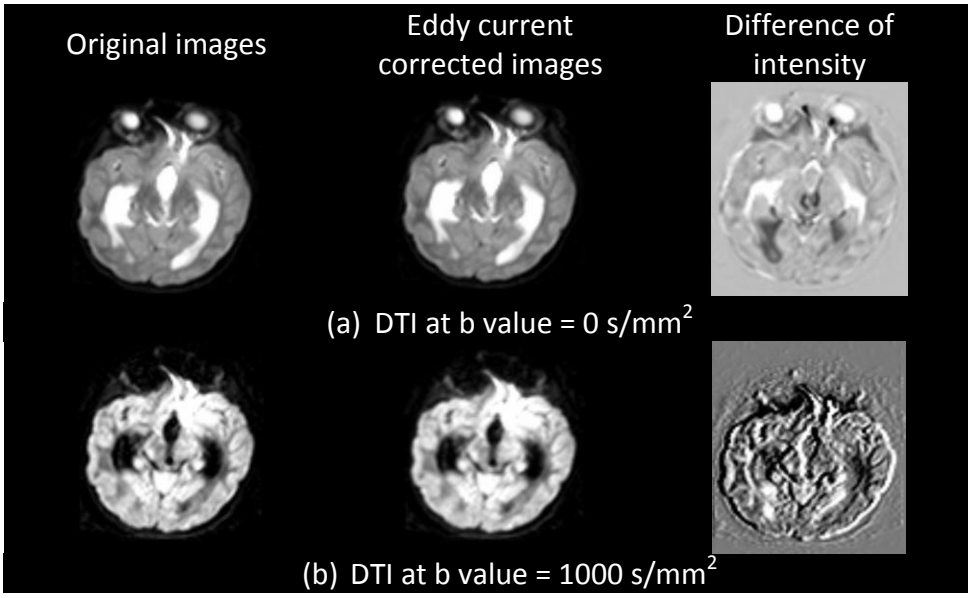


Figure 5.1: A comparison between original and eddy current corrected images. Eddy current correction applied on DTI data at b-value = 0 s/mm^2, and b-value = 1000 s/mm^2. The difference of image intensity is a subtraction of eddy current corrected image from the original image (case MB532).

5.3.2 Skull Stripping

From the observation of skull stripping results, threshold values can be set to 0.3-0.35 for MRI and to 0.25 for diffusion MRI. On gross dataset, the BET method can remove most of non-brain tissues on T2, FLAIR, pre-contrast T1, post-contrast T1, DWI and DTI, as shown in Figure 5.2. According to FSL's guideline, the results are acceptable although some non-brain tissues such as eyeballs and nasal bone are not entirely removed. However, the result is deemed unacceptable if brain tissue is removed after skull stripping process.

The BET method occasionally performed poorly on some dataset as shown in Figure 5.3. The brain extraction produced hollow region within brain tissue on DWI of Figure 5.3 (d) and especially on FLAIR images of Figure 5.3 (a) and (b). Although the threshold was set to a very low value, the brain tissue was not properly extracted. In the case of post-contrast

T1 images of Figure 5.3 (c), partial brain tissue was removed due to the similar intensity of brain tissue and the non-brain tissues. In the cases of poor skull stripping, the original brain imaging data are used instead for image registration.

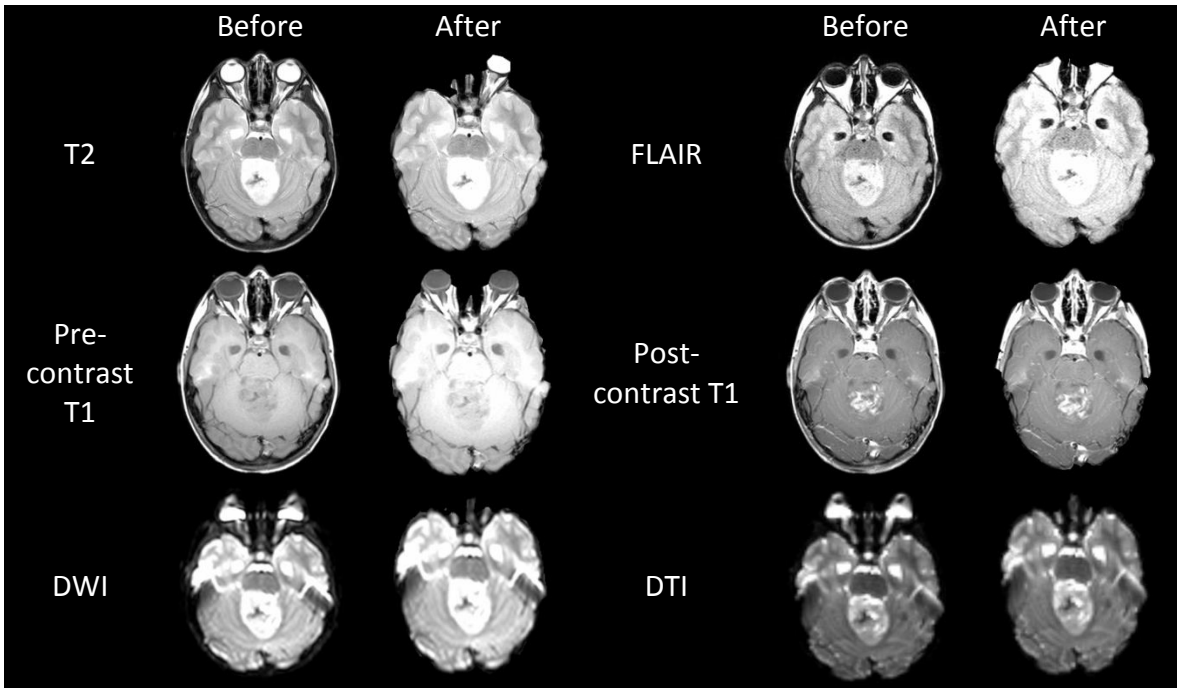


Figure 5.2: Satisfactory skull stripping outcomes obtained from BET method. Before and after skull stripping of a medulloblastoma based on T2, FLAIR, pre-contrast T1, post-contrast T1, DWI and DTI (case MB783).

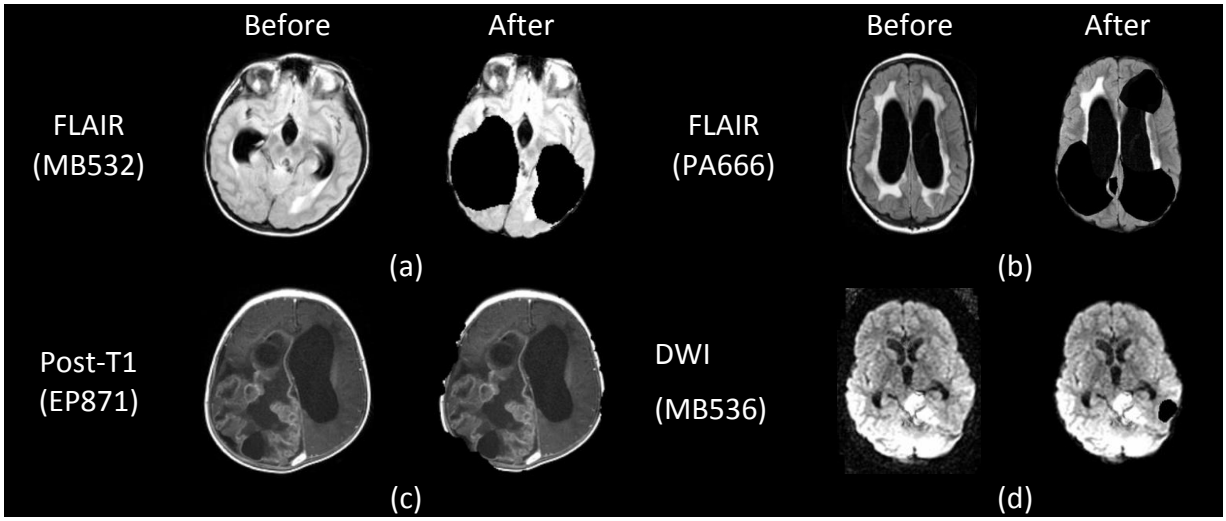


Figure 5.3: Failure of skull removal on some datasets. The BET method produced hollow regions within the brain tissue on some types of MRI: FLAIR, post-contrast T1 and DWI.

5.3.3 Diffusion MRI Reconstruction

Diffusion MRI reconstruction is conducted with the use of standard equations, explained in section 4.2.3, and functions in FSL. Examples of ADC, MD and FA maps reconstructed from DTI are shown in Figure 5.4. Observing the image characteristics, a FA map has distinctive features compared to a MD and ADC map, whereas the image features of MD and ADC map is similar. This is because FA aims to measure value of anisotropic diffusion, whereas ADC and MD attempt to measure an average of diffusion coefficient in each imaging voxel.

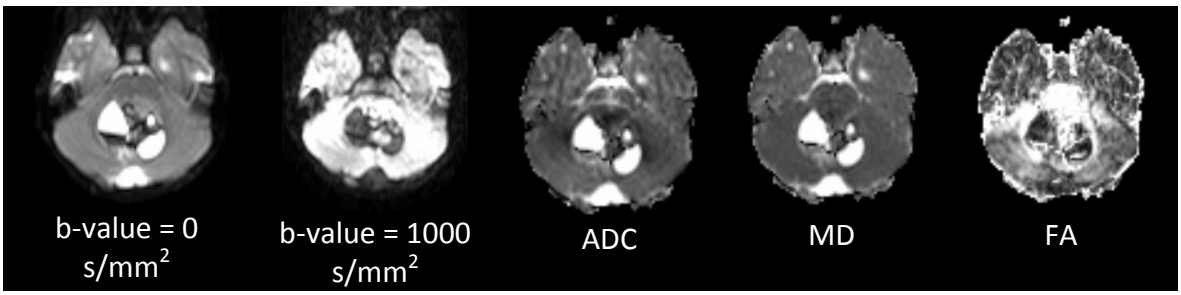


Figure 5.4: Reconstruction of DTI images produces ADC, MD and FA maps (case MB718).

5.3.4 Image Registration

After the whole brain extraction was conducted before registration, the floating images: pre-contrast T1, post-contrast T1, FLAIR, ADC, MD and FA images are registered on the T2 reference images within the same subject. Based on FLIRT method, correlation ratio (the default similar measure) was compared against normalised mutual information (cost function). Based on visual inspection, FLIRT method with the correlation ratio produces corrupted registration results for some cases, whereas FLIRT method with normalised mutual information produces a more satisfactory outcome on the same cases. For

example, post-contrast T1 images are not registered on T2 images correctly on three study cases, as shown in Figure 5.5.

The registration with normalised mutual information worked well for all cases and all image types used in this thesis. For example, linear registration for a case involving a medulloblastoma tumour is shown in Figure 5.6. In addition, the interpolation based on sinc function is recommended over the interpolation based on trilinear (120). As a result, FLIRT with 6 DOF, normalised mutual information of cost function and sinc interpolation is used throughout the thesis.

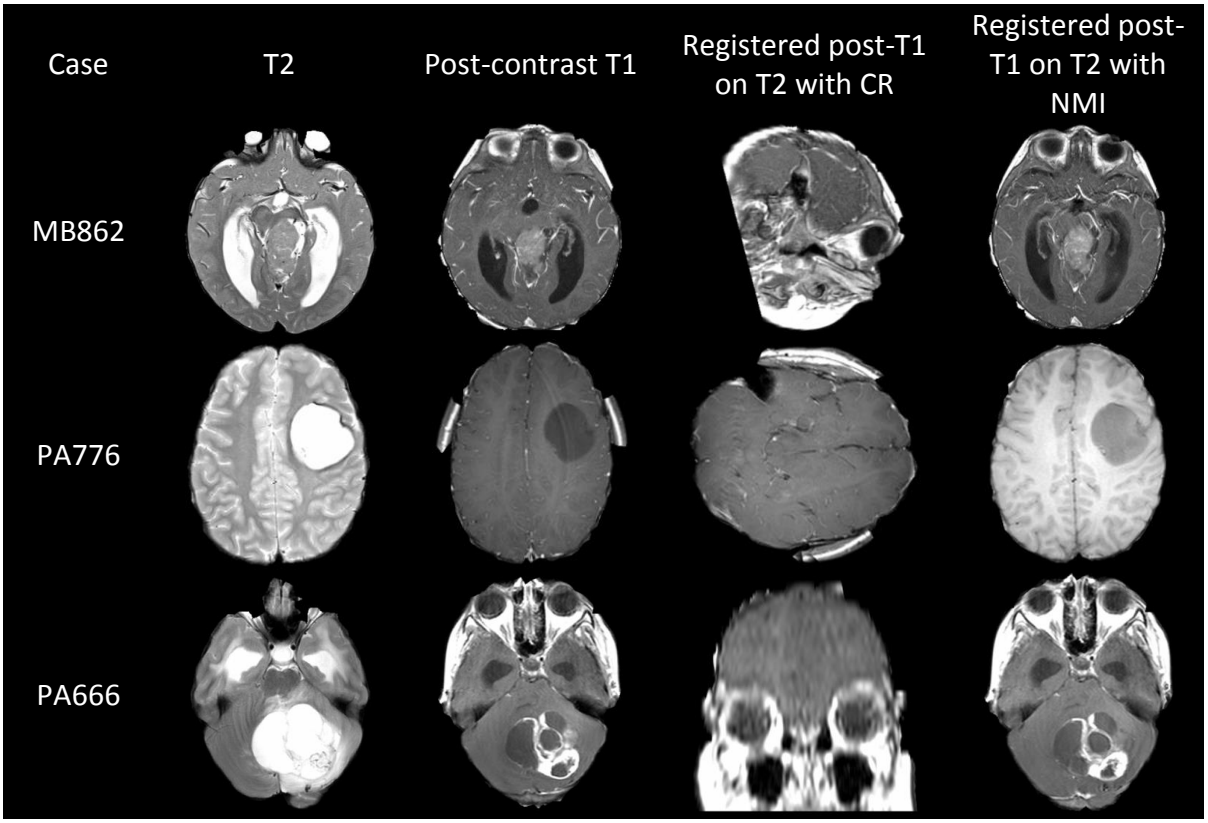


Figure 5.5: Failure of registration based on CR similarity measure. Cases that are registered incorrectly with CR function but registered correctly with NMI function.

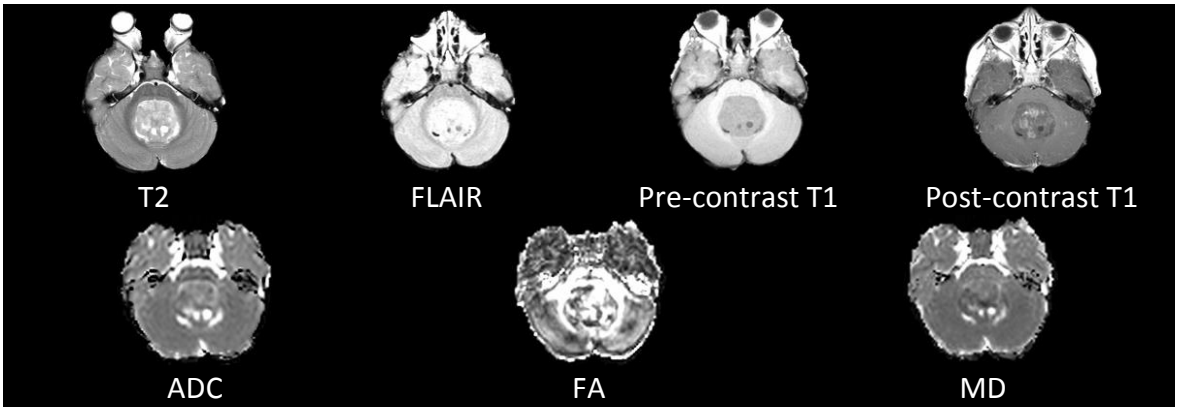


Figure 5.6: Successful registration for a medulloblastoma with NMI cost function shows the same slice of T2, FLAIR, pre-contrast T1, post-contrast T1, ADC, FA and MD (case MB719).

5.3.5 Intensity Normalisation

The original images, images normalised by using the decile-based piecewise linear transform method and images normalised by using the linear normalisation method are displayed in Figure 5.7. For all three tumour types and all MR image types, the images normalised by using the linear normalisation method are brighter and have higher contrast than the original images and images normalised by using the decile-based piecewise linear transform method. This is because linear stretching provides a wider range of intensity transformation and enhances image contrast, thus better highlighting the pathological region.

The intensity normalisation methods were examined by comparing sum of mean difference square, SMD between different tumour types (Table 5.3). The comparison of intensity normalisation methods shows that the linear normalisation method provides higher mean intensity difference between different tumour types, for most of MRI types, except for pre-contrast T1-weighted images. This implies that the linear normalisation

method provides the better classification of pathological tissues than the original image intensity and the decile-based piecewise linear transform method. As a result, the linear normalisation method is selected in this thesis.

It should be noted that the two methods evaluated in this thesis are standard techniques widely used in MRI pre-processing. It is beyond the scope of this work to perform a comprehensive review of intensity normalisation techniques. However, for potential future studies the technique proposed by Ekin (292), presenting intensity normalisation method for MR images of pathological tissues, would be a useful refernce.

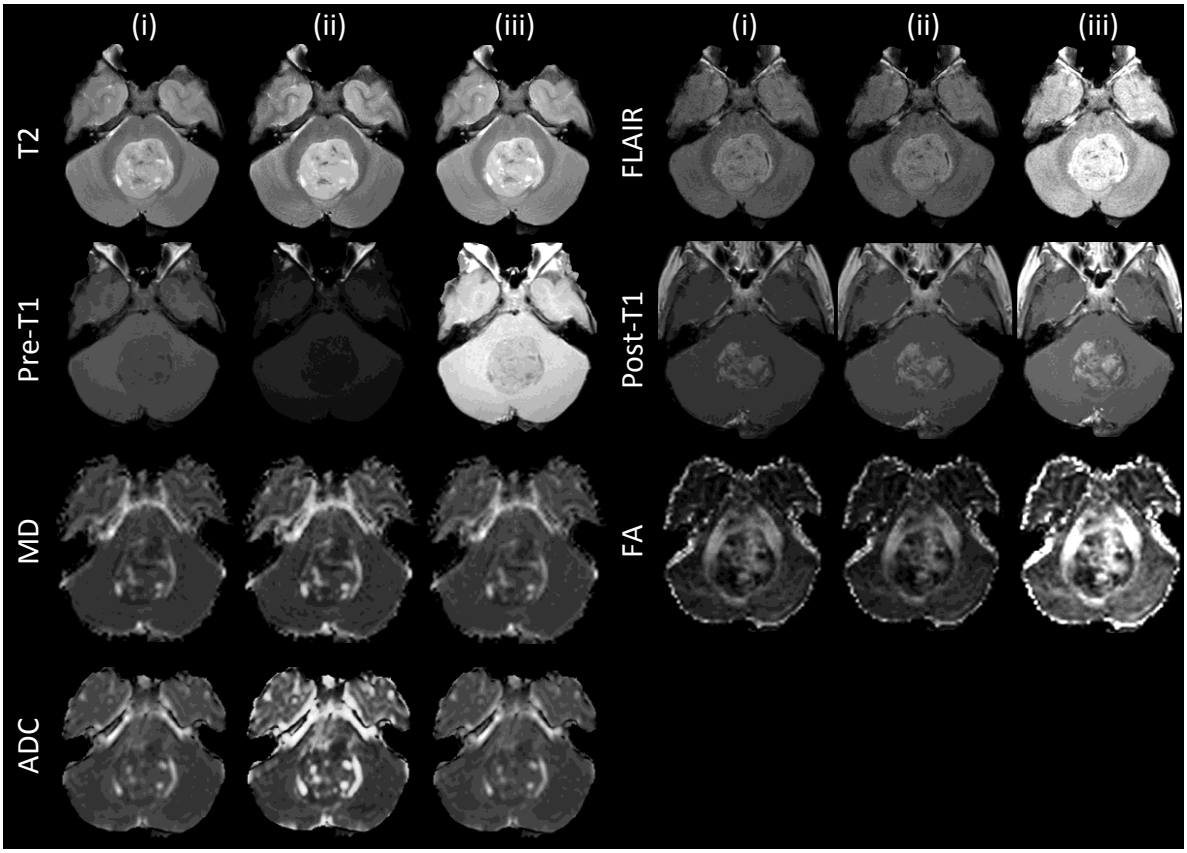


Figure 5.7: A comparison of intensity normalisation performed on a medulloblastoma (case MB783). For each MRI type: (i) original, (ii) applied decile-based piecewise linear transform and (iii) applied linear normalisation images.

Table 5.3: Mean and standard deviation of intensity for each tumour and MRI type. The value of sum of mean difference square, SMD summarises the power of discrimination of the intensity normalisation methods.

	ADC	MD	FA	T2	Pre-T1	Post-T1	FLAIR
				Original Image			
MB	-0.12(0.00)	-0.86(0.51)	0.59(1.04)	-0.12(0.91)	0.05(0.94)	-0.02(0.91)	-0.22(0.55)
PA	0.15(1.52)	0.77(0.83)	-0.51(0.66)	0.12(0.97)	-0.05(1.14)	0.04(1.09)	0.19(1.38)
EP	-0.11(0.01)	0.50(0.49)	-0.39(0.76)	-0.04(1.20)	0.01(0.91)	-0.04(0.99)	0.09(0.80)
SMD	0.05	1.52	0.72	0.03	0.01	0.00	0.09
				Decile-based Piecewise Linear Transform			
MB	-0.35(0.33)	-0.81(0.49)	0.65(0.89)	-0.63(0.37)	0.16(0.92)	-0.05(0.84)	-0.15(0.53)
PA	0.13(1.03)	0.71(0.83)	-0.59(0.68)	0.57(1.10)	-0.40(1.01)	0.14(1.21)	0.12(1.06)
EP	0.26(1.41)	0.50(0.82)	-0.36(0.94)	-0.18(0.82)	0.41(0.92)	-0.19(0.77)	0.07(1.61)
SMD	0.21	1.36	0.86	0.73	**0.34**	0.06	0.04
				Linear Normalisation			
MB	-1.06(0.38)	-0.94(0.43)	0.70(0.83)	-0.75(0.74)	0.11(0.80)	-0.05(1.07)	-0.04(0.75)
PA	0.72(0.70)	0.86(0.64)	-0.73(0.60)	0.63(0.80)	-0.14(1.17)	0.17(0.99)	0.17(1.05)
EP	0.15(0.77)	0.51(0.52)	-0.21(0.92)	-0.12(0.89)	0.06(0.99)	-0.35(0.89)	-0.27(1.37)
SMD	**1.66**	**1.82**	**1.05**	**0.96**	0.03	**0.15**	0.10

5.4 Conclusions

This chapter presented the procedures for pre-processing before applying pattern recognition methods, described in the subsequent chapters. The process includes artefact removal, skull stripping, diffusion MRI reconstruction, registration, and intensity normalisation. From this study, it is learnt that default parameters recommended in the standard tool do not always work well for some cases in our study; therefore an experiment was conducted to select the optimal parameters for our dataset.

Skull stripping is recommended as aforementioned in section 3.2.2 to enhance registration performance. It is used as a part of the intensity normalisation method to scale intensity within the brain tissue region. However, it was found that skull stripping does not work well on some cases. Therefore registration was done with the original data for those cases. In registration, the default parameters of FLIRT did not perform well with our dataset. We found that normalised mutual information gives a better result than correlation ratio; hence normalised mutual information is used throughout the thesis.

In order to standardise image intensity, it was initially reviewed that decile-based piecewise linear transform method is the ease of use and provides a suitable image intensity range for classification. However, empirically linear normalisation, a simpler technique, produced higher sum of mean difference square between tumour types. This implies that using the linear normalisation method should result in better classification accuracy. As a result, it is selected for the work in this thesis.

CHAPTER 6 EXPERIMENT 2: BRAIN TUMOUR SEGMENTATION

This chapter presents the hybrid segmentation and evaluation methods, addressing the first major objective of section 1.2. The methodology of experiment 2 is described in section 4.3. The hybrid segmentation method involving graph and active contour based image segmentation method is proposed and evaluated against the gold standard.

6.1 Introduction

In medical image analysis, the characterisation of histological types of brain tumours begins with the definition of a region of interest using a segmentation method. A segmentation method is used to extract target from irrelevant regions. In experiment 2, the goal is to extract brain tumour regions from healthy tissues.

The hybrid semi-automatic segmentation based on the MNcut and GVF snake methods, as reviewed in section 3.3, is hypothesised to improve the capture range of brain tumour boundary. The method is applied on four types of MRI: T2, FLAIR, pre-contrast T1 and post-contrast T1. Subsequently four evaluators provide rating scores for tumour outlining on the four MRI types. A single type of MRI obtained the highest scores are selected and evaluated by comparing with manual segmentation.

6.2 Material

The most common paediatric brain tumour types with pre-treatment MRI were considered; pilocytic astrocytoma, medulloblastoma, and ependymoma. The acquisition of T2, FLAIR, pre-contrast T1, and post-contrast T1 images was conducted in three separate centres using 1.5T GE, 1.5T and 3T Phillips scanners, following a common protocol defined by the CCLG Functional Imaging Group. All images were stored in the CCLG Functional Imaging Database. The number of PA, MB, and EP cases available at the time this study was undertaken is 15, 14 and 8 respectively. The average and range of TR, TE, TI, resolution and slice thickness of the brain imaging data are shown in Table 6.1. The study cases with value of TR, TE, TI, resolution, FOV, slice thickness and magnetic field strength as well as scanner's manufacture used in experiment 2 are listed in section C.2, Appendix C.

Table 6.1: Imaging parameters for experiment 2. Average and range of TR, TE, resolution and slice thickness are presented in the format of average (min, max).

Image type	TR (ms)	TE (ms)	Resolution (mm)	Slice thickness (mm)
T2	4617 (2660,9140)	89 (77,105)	[11.3,12.6,5.8] ([0.4,0.4,3.9], [402,448,50])	3.8 (3, 5)
FLAIR	10071 (8002,11000)	119 (89,140)	[0.6,0.6,5.1] ([0.4,0.4,4.4], [1,1,6.6])	4.4 (4, 6)
Pre-T1	610 (360,1400)	13 (8,18)	0.6,0.6,5.1] ([0.4,0.4,4.4], [0.9,0.9,6.5])	4.3 (4, 5)
Post-T1	548 (250,828)	15 (2.3,22)	[0.6,2.5,4.9] ([0.4,0.4,4.4], [0.9,0.9,6.6])	4.1 (3, 6)

TI of FLAIR = 2511 (2000, 2800) ms.

6.3 Results and Discussion

This content presented in this section is organised using the following sub-sections: 1) segmentation outcomes, 2) evaluation based on scoring system, 3) evaluation based on supervised method, 4) constraints of the proposed method, 5) adaptation to the clinical environment and 6) future work.

6.3.1 Segmentation Outcomes

The segmentation results for T2, FLAIR, pre-contrast T1 and post-contrast T1 obtained from the proposed semi-automatic segmentation method are shown in Figure 6.1. For example in Figure 6.1 (a), a T2 image is cropped to reduce computational load and then initially segmented by using the MNcut method. The MNcut method produces multiple segments. Seven sub-regions within the T2 segmented images were selected to form an entire tumour region where the boundary was drawn. The initial boundary was subsequently adjusted by using the GVF snake method. The gradient of edge map and GVF field map produced by the GVF snake method are computed. The final boundary is adjusted by using the GVF snake method, showing the improved capture of the abnormal region. The segmentation results depend on parameter values of the semi-automatic segmentation method. For example, varying K value of the MNcut method, while other parameters of the GVF snake method are fixed, the segmentation based on FLAIR image in Figure 6.2 can yield different tumour outlines. A higher K value generally produces a better initial boundary, although it may require a higher level of manual input and longer computation time.

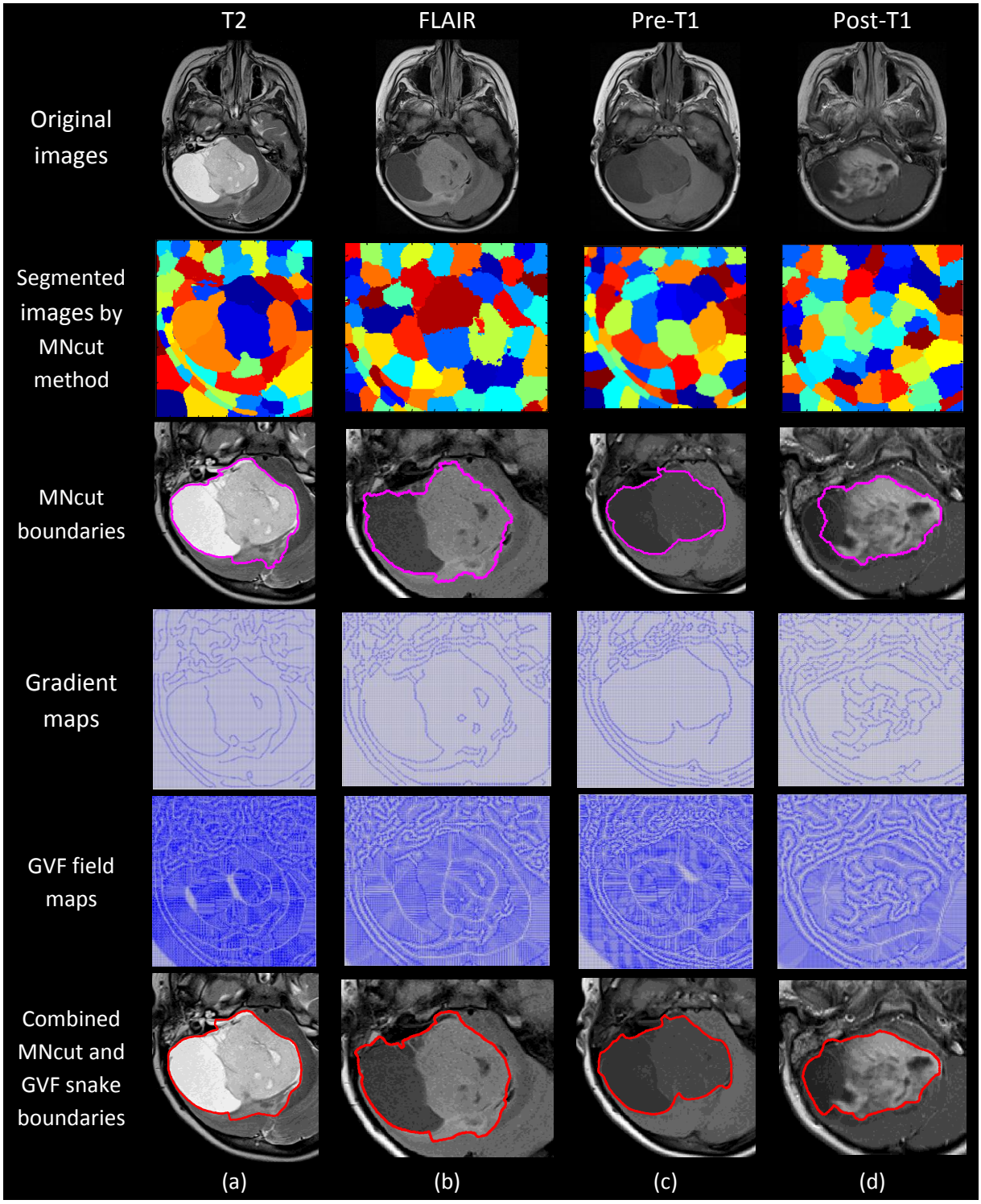


Figure 6.1: The four MRI types of a pilocytic astrocytoma are segmented by the combined MNcut and GVF snake method. Tumours are initially segmented by the MNcut method and the derived tumour boundaries are refined by the GVF snake method (case PA873).

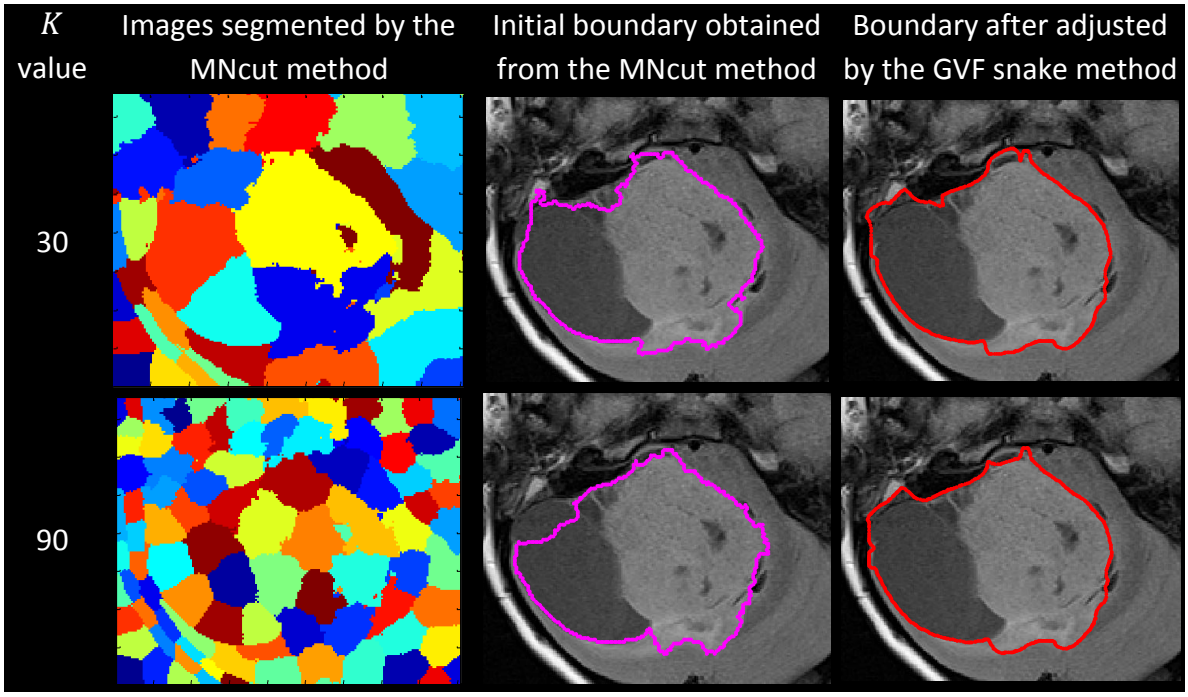


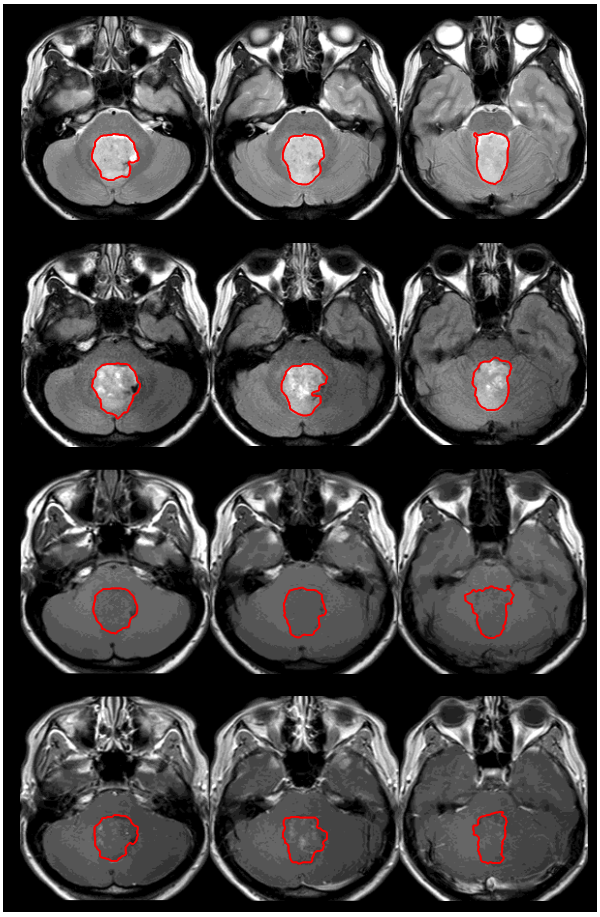
Figure 6.2: Segmentation results derived from two K values of the MNcut method. FLAIR image based segmentation with different K values of the MNcut method, and fixed parameters of GVF snake method (σ=1, threshold of Canny detector = 0.2) (case PA873).

For the MNcut method, the K value is initialised to 60. A variation is applied if the segmentation result is not satisfactory. Experimentally, K values in the range [30, 150] are selected depending on the image type by adjusting the value in steps of 10. Compared to other MR image types, T2 images have better image contrast and generally require lower K values. The suitable number of segments was experimentally determined to be between 30-100 segments for T2 images. The mid-range either 60 or 80 number of segments is more commonly used.

For GVF snake method, according to He et al. (290)'s study, the parameters are set as follows. The Gaussian filter is used to blur images with σ set to between 1 to 3. The Canny edge detector is used to find an edge map with threshold set between 0.1 to 0.4, depending on tumour structure, μ = 0.2, α = 1 and β = 1.

6.3.2 Five-Point Likert Scale of Evaluation

Brain tumour boundaries extracted from four types of MRI (T2, FLAIR, pre-contrast T1 and post-contrast T1) and from three tumour types (medulloblastoma, pilocytic astrocytoma and ependymoma) were assessed by using the five-point Likert scale of evaluation criteria. The evaluation was performed using visual inspection by four trained observers with knowledge of brain anatomy and tumour morphology. The evaluators rated the quality of semi-automatic segmentation results based on the five-point Likert scale of evaluation criteria, as demonstrated in Figure 6.3.



T2

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
5	4	3	2	1

FLAIR

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
5	4	3	2	1

Pre-contrast T1

Strongly Agree	Agree	Neutral	Disagree	Strongly disagree
5	4	3	2	1

Post-contrast T1

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
5	4	3	2	1

Figure 6.3: A segmentation evaluation form. The tumour boundaries are semi-automatically segmented from T2, FLAIR, pre-contrast T1 and post-contrast T1 (MB796).

The boundaries drawn for each MRI type varies in relation to a visible abnormal region (Figure 6.4). The boundaries extracted from T2, FLAIR and pre-contrast T1 images generally reveal the capture of a larger tumour region because these image types have greater sensitivity to pathology. The boundaries obtained from post-contrast T1 images tend to have a smaller tumour region because these images have lesser extent of visible tumour, their enhanced patterns are more heterogeneous and tumour boundaries are less well defined.

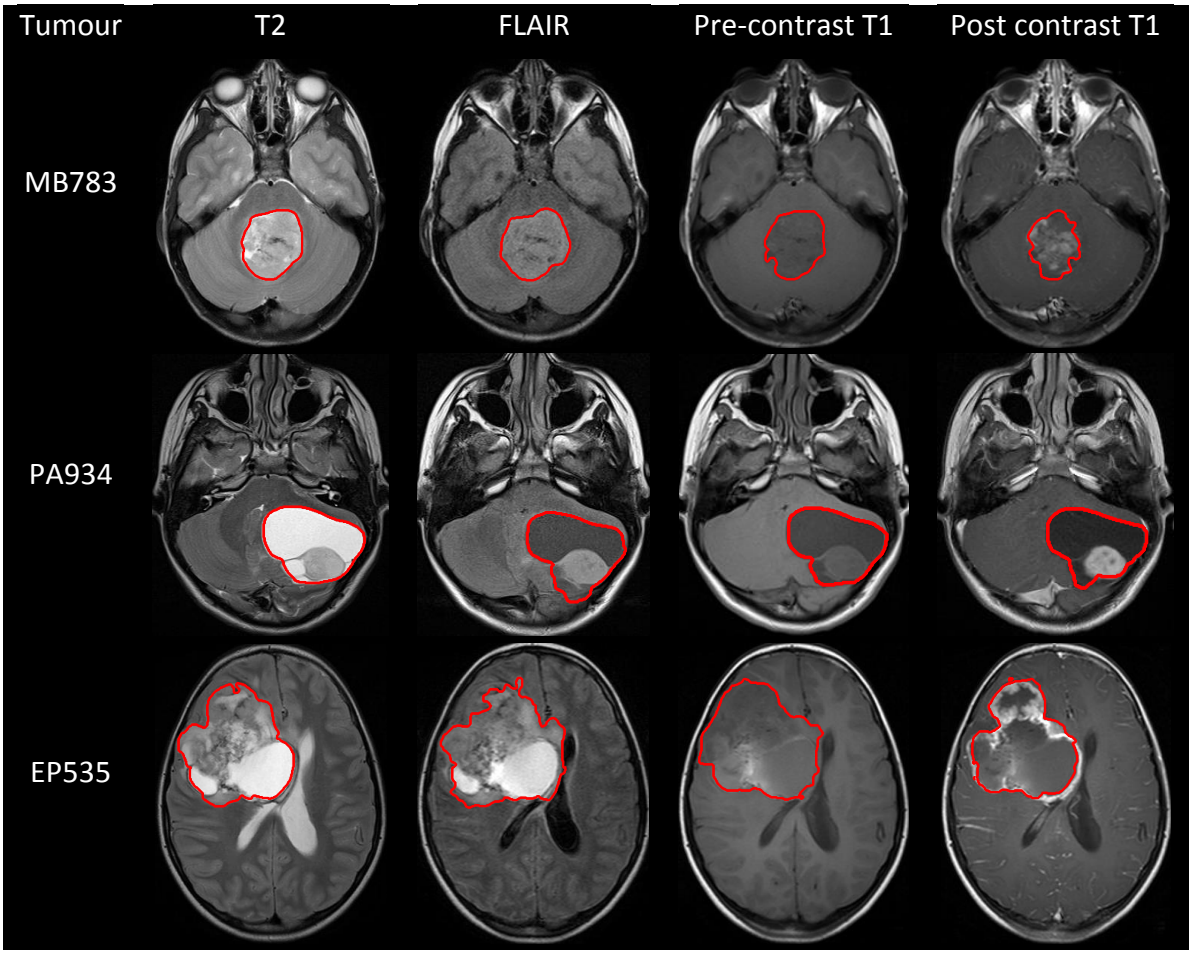


Figure 6.4: Semi-automatic segments of medulloblastoma, pilocytic astrocytoma, and ependymoma based on T2, FLAIR, pre-contrast T1, and post-contrast T1.

From the five-point Likert scale of evaluation, the average score rated by the four evaluators and the inter-rater reliability on the ratings is shown in Figure 6.5. On average, the evaluators rated the T2 based segmentation as highest, while post-contrast T1 based segmentation was rated as lowest (Figure 6.5 (a) – (d)). However, one of the evaluators rated the T2 based segmentations on pilocytic astrocytomas slightly lower compared to the FLAIR based segmentation (Figure 6.5 (c)). On average, the inter-rater reliability of the T2 and post-contrast T1 based segmentation are over 0.8 (Figure 6.5 (e)).

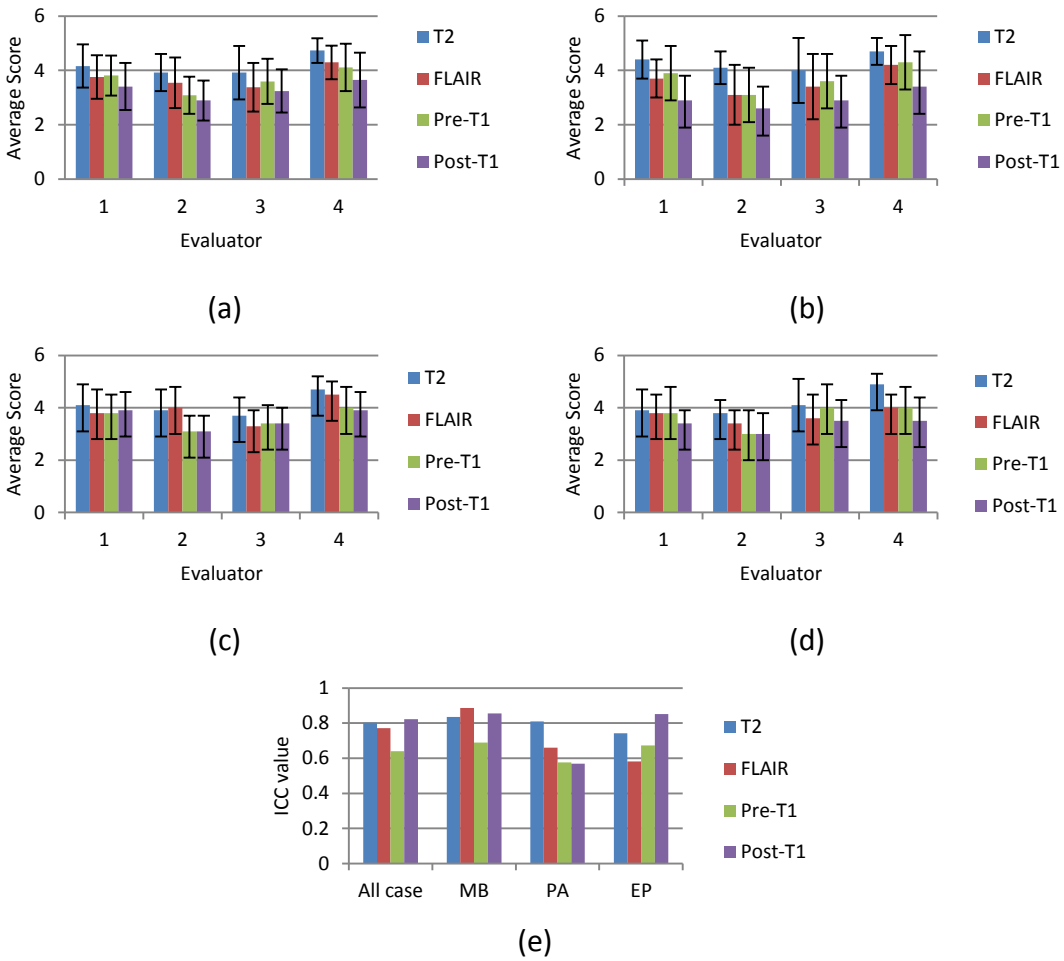


Figure 6.5: The average score with standard deviation obtained from segmentation scores rated by the four evaluators (a) all tumour types, (b) medulloblastoma cases, (c) pilocytic astrocytoma cases, (d) ependymoma cases and (e) intraclass correlation coefficient. The average score and standard deviation are indicated by a bar height and error bar.

The scores of the FLAIR and pre-contrast T1 based segmentation are relatively similar and generally have lower scores than the T2 based segmentation (Figure 6.5 (a) – (d)). The inter-reliability of rating FLAIR and pre-contrast T1 based segmentation varied depending on tumour types (Figure 6.5 (e)).

From the evaluation based on five-point Likert scale, the proposed semi-automatic segmentation method performs better for some image types, particularly for T2 images because its sensitivity to lesions and oedema provides greater visibility of pathological tissues and hence offers higher distinction between tissues. The method performs less well for post-contrast T1 images than the other image types because they appear more heterogeneous and consecutive homogeneous regions are difficult to define.

Improved segmentation results can be achieved with parameters optimised for each image type, particularly for those with greater heterogeneity of post-contrast T1. In some cases, the contrast agent enhances pathological tissues and exhibits a thick rim-enhancing lesion. The semi-automatic segmentation method can sometimes only capture the inner boundary of a rim-enhancing lesion on post-contrast T1 images because the MNcut method separate two regions of the edges and the GVF snake method can only capture one part of the edges.

The evaluation based on Likert scale criteria is known as a subjective evaluation, which may have two factors influencing the evaluators' scores. One is that how well the delineation of the semi-automatic method matched the visible abnormality on the image. Another is that how well the delineation of the semi-automatic method matched the tumour border as judged from all of the image types together. The first of these

corresponds to the performance of the semi-automatic segmentation method itself for the image, while the second includes an additional evaluation of the ability of that image type to capture the entire extent of the tumour border. The amount of emphasis placed on each of these factors will probably have been different for each observer.

Despite these limitations of the subjective evaluation, the optimal delineation of brain tumour region produced by this segmentation method was rated to be based on T2-weighted images, which corresponds to the study of Tsuchiya et al. (94); stating that T2-weighted images are the most sensitive for detecting intracranial lesions. In addition, the results of the subjective evaluation showed that the level of importance of the image type is in agreement with the way MR image types used in a clinical delineation of a brain tumour in children.

Clinically, the delineation of a brain tumour is performed on a single image type, but with information gathered from examining a combination of MR sequences. Lesley Macpherson, a radiologist at University Hospitals Birmingham NHS Foundation Trust, gave her viewpoint on deciding brain tumour boundary in children; it highly depends on T2, moderately depends on FLAIR and pre-contrast T1, and occasionally depends on post-contrast T1 images. The tumour boundary is generally delineated on T2 images (Macpherson L., personal communication, October 22, 2013). However, Nigel P. Davies, a lead MRI Physicist at University Hospitals Birmingham NHS Foundation Trust, pointed out the difference of outlining adult brain tumours. Outlining adult high-grade tumours is more dependent on post-contrast T1 images, while outlining adult low-grade tumours is more commonly done on T2 and/or FLAIR images (Nigel P. Davies, personal

communication, July 4, 2014). As a result, the evaluation outcome of the semi-automatic segmentation method can be different for outlining adult brain tumours.

6.3.3 Supervised Evaluation

From the five-point Likert scale of evaluation results, the semi-automatic segmentation can delineate the boundary of abnormal tissue on T2 better than FLAIR, pre-contrast T1, and post-contrast T1 images. To further quantify the quality of the proposed segmentation method based on T2 images, manual segmentation was conducted. The semi-automatic segmentation and manual segmentation was conducted by the same person, who has knowledge of brain anatomy and tumour morphology, because the decision of outlining tumour border should be consistent between the two segmentation approaches. An example of brain tumour outlining obtained from the manual and semi-automatic segmentation method is shown in Figure 6.6.

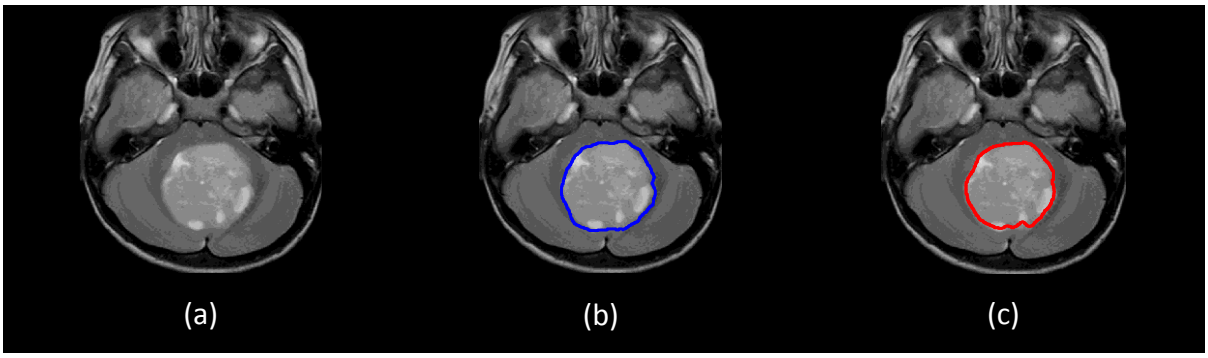


Figure 6.6: A comparison of tumour boundaries derived from manual segmentation and semi-automatic segmentation based on T2 image (a) original image, (b) tumour boundary obtained from manual segmentation and (c) tumour boundary obtained from the semi-automatic segmentation method (case MB532).

A quantitative comparison was performed based on the average and standard deviation of sensitivity, specificity, and modified Hausdorff distance (MHD) and volume overlap

ratios and dice coefficient, as shown in Table 6.2. High sensitivity, high specificity, and short distance computed by using MHD imply high similarity between manually and semi-automatically segmented tumour volumes. On average, the volume overlap ratios show that the semi-automatic and manual segmentation capture a similar region, with mean of the volume overlap over semi-automatically segmented volume (Ratio1), volume overlap over manually segmented volume (Ratio2) and dice coefficient being 95%, 90%, and 93% respectively.

Table 6.2: Comparison between manual segmentation and semi-automatic segmentation. The evaluation is based on sensitivity, specificity and modified Hausdorff distance, ratio 1, ratio 2 and dice coefficient. Average (standard deviation) of the evaluation criteria is represented.

Evaluation	MB	PA	EP	Total
Sensitivity	91.0% (3.1)	90.5% (6.1)	89.2% (3.4)	90.4% (4.6)
Specificity	99.4% (0.6)	99.3% (0.8)	99.0% (0.7)	99.3% (0.7)
MHD	0.8 (0.3)	0.9 (0.5)	1.1 (0.5)	0.9 (0.4)
Ratio 1	97.1% (2.6)	93.6% (8.0)	95.0% (4.0)	95.2% (5.8)
Ratio 2	91.0% (3.1)	90.5% (6.1)	89.2% (3.4)	90.4% (4.6)
Dice coefficient	93.8% (1.6)	91.6% (4.5)	91.8% (1.8)	92.5% (3.3)

Compared to manual segmentation, semi-automatic segmentation of medulloblastoma tends to have more similarity and less deviation than that of pilocytic astrocytoma and ependymoma. This is because simpler pathological structures of medulloblastoma are relatively easier to delineate. The Ratio 1 shows higher value than the Ratio 2 and the dice coefficient. This implies that the semi-automatically segmented volumes could be under-segmented or have smaller volumes in comparison with manually segmented volumes.

The non-overlapping volumes are the volumes which may be misclassified by either the manual or semi-automatic segmentation method.

Experiment 2 is conducted on a computer with the specification of CPU intel Core i5 @ 2.4 GHz and RAM 6 GB. Based on this machine, the computation time for semi-automatic segmentation method of each slice is approximately 30-120 seconds, while the time required for manual segmentation depends on the level of expertise of the operator.

6.3.4 Constraints of the Proposed Method

The proposed segmentation method provides extraction of abnormal brain region based on T2 better than FLAIR, pre-contrast, post-contrast T1 (see Figure 6.5). The method offers better coverage of abnormal region than using a single technique alone based on visual inspection (see Figure 6.1). The MNcut method produces an initial boundary which reduces the number of seed points for the GVF snake technique and the chance of the snake contour moving towards the wrong boundary. The GVF snake method refines the boundary obtained from the MNcut method and offers an optimal outlining. Furthermore, the supervised evaluation results showed over 90% similarity between the boundaries obtained from manual and semi-automatic segmentation methods (see Table 6.2). As a result, this segmentation method is suitable for the purpose of our study.

Although the proposed hybrid semi-automatic segmentation method offers several advantages, some limitations of this approach need to be improved. The main constraint of this segmentation technique is the requirement of two main interactions from the operator; the input number of segments for the MNcut method and the setting of the

threshold for edge detection used in the GVF snake technique. These parameters rely heavily on the tissue structures and image contrast.

A default number of segments can be initialised so that there is no requirement from the operator; however, the number of segments is data dependent and an appropriate value can be manually selected to optimise the result. The higher the number of segments, the higher the number of patterns obtained. A tumour image having less complex structures or a well-defined tumour boundary generally requires a smaller number of segments.

In edge detection used in the GVF snake technique, higher threshold values may be required to remove lines either within or surrounding the abnormal region to avoid tracing toward an unsatisfactory boundary. On the other hand, lower threshold values may be preferred to maintain the boundary of a faint contour.

6.3.5 Adaptation to Clinical Environment

The proposed technique is simple, reduces manual input, offers faster operation and produces high accuracy compared to manual segmentation. Due to the ability of the GVF snake algorithm to progress towards the tumour boundary, an operator only needs to select certain sub-regions produced by the MNcut method that make up the ROI. As the required manual input to GVF snake is the boundary of a composite region, the operator can select any coordinate within a sub-region in order to mark it as selected. This is much less labour-intensive and the operator does not require extensive oncological expertise to perform this task. The resulting ROI is then used in the next steps of the processing pipeline to produce an automated diagnosis. This could be advantageous in a clinical setting to complement conventional radiological diagnosis.

6.3.6 Future Work

Although the combined MNcut and GVF snake method provides the optimal segmentation for abnormal brain region based on T2 datasets, it still requires further improvement to provide differentiation of distinct tissue compositions within the abnormal region. Single modality MRI cannot provide sufficient information for extracting sub-regions. Post-contrast T1-weighted images demonstrate blood-brain barrier breakdown and highlight specific tissues such as necrosis and active tumour region (293). FLAIR images represent more vivid depiction of cystic region and better differentiate between maldevelopmental/porencephalic and neoplastic/inflammatory lesions than do other MRI (294). FLAIR can also distinguish between cystic intracranial lesions and CSF with high sensitivity (88). As a result, the use of multimodal MRI analysis can provide added advantages of differentiating tissue types.

Because T2-weighted image based segmentation empirically provides better tumour delineation, T2-weighted images can potentially be used as reference images for registration and consequently segmentation of other MR image types. Within the T2 based segmented region, the other co-registered images can be segmented by using the MNcut method. The sub-regions obtained from each image type could be used to identify distinct components within abnormal tissues.

6.4 Conclusions

This chapter presented a hybrid semi-automatic segmentation technique, based on the MNcut and GVF snake methods. The method was applied to four types of structural MRI:

T2, FLAIR, pre-contrast T1, and post-contrast T1 and three conclusions were drawn as follows.

First, this approach improves upon the MNcut or GVF snake method alone by increasing the capture range of tumour boundary delineation. The MNcut method provides an initial contour to reduce the number of input parameters for the GVF snake method and prevents the snake falling into local minimum (or progress into the wrong edges). The GVF snake method subsequently adjusts jagged contours produced from the MNcut method and extends the capture range of abnormal region.

Second, the five-point Likert scale evaluation revealed that the T2-based semi-automatic segmentation method provided a better delineation of tumour boundary in comparison with segmentation based on other image types. In addition, the tumour boundary obtained from the T2-based semi-automatic segmentation method has relatively similar capture range to the manual segmentation approach.

Third, the proposed hybrid semi-automatic segmentation technique promises to be useful in clinical settings, because it can be performed within a shorter timeframe than manual segmentation and does not require extensive neuroanatomical expertise. It could also be used as the first step in a manual segmentation allowing rapid refinement.

In terms of future work, to separate abnormal components, such as cystic regions from solid tumours, the combination of multi-sequence MR image based segmentation should be further investigated.

CHAPTER 7 EXPERIMENT 3: CLASSIFICATION OF BRAIN TUMOURS BASED ON TEXTURE ANALYSIS OF SINGLE MODALITY MRI

This chapter presents the methods of pattern recognition employed to differentiate childhood brain tumours based on textural features of single MRI modality, supporting the second objective of section 1.2. This process includes TA for feature generation, feature selection for dimensionality reduction and, classification for the discrimination of tumour types. Classification performance derived from the TA of individual MR modality is validated and compared using several validation methods and statistical hypothesis testing.

7.1 Introduction

Textural features derived from MR images have been shown to provide high sensitivity for the characterisation of histological types of brain tumours. In order to obtain a robust textural feature set, an appropriate feature selection method is required to reduce dimensionality, to define a good feature combination and so to enhance classification rate, as reviewed in section 3.6.

This chapter investigates the two different feature selection techniques: PCA and the combined mRMR and SFS method. PCA is a popular feature extraction and dimensionality reduction method, which has been used for MR image based brain tumour classification

(19, 258). Mutual information based feature selection has also been widely applied on medical data analysis (259-261). Minimum redundancy and maximum relevance proposed by Peng et al. (257), is one of popular MI based feature selection methods, which has been useful for classifying MR image based brain diseases including brain tumours (262-264).

A better feature selection method obtained from the experiment is subsequently used in the classification process to identify which image type can provide better discriminative power through the use of texture analysis. The theoretical background of these feature selection methods were described in section 3.6. The design of experiment 3 is illustrated in the flowchart (Figure 4.4) in section 4.4.

7.2 Material

Three types of pre-treatment paediatric brain tumours have been considered: embryonal, astrocytic and ependymal tumours. The number of embryonal, astrocytic and ependymal cases available at the time this study was undertaken are 12, 8 and 6 cases. The number of subtype is shown in Table 7.1. All cases were obtained from the CCLG database. The structural and diffusion MRI scans were acquired from four centres using a 1.5T GE, 1.5T Siemens, 1.5T and 3T Phillips scanners, following a common protocol defined by the CCLG Functional Imaging Group. The average and range of TR, TE, TI, resolution and slice thickness of the brain imaging data are shown in Table 7.2. The study cases with value of TR, TE, TI, resolution, FOV, slice thickness and magnetic field strength as well as the scanners used in experiment 3 are listed in section C.3, Appendix C.

Table 7.1: Number of subjects for each tumour type for experiment 3.

WHO Defined Tumour	Embryonal		Astrocytic	Ependymal	
Subtype	medulloblastoma	Desmoplastic /nodular medulloblastoma	Pilocytic Astrocytoma	Ependymoma	Anaplastic Ependymoma
Number of cases	11	1	8	2	4

Table 7.2: Imaging parameters for experiment 3. Average and range of TR, TE, resolution and slice thickness are presented in the format of average (min, max).

Image type	TR (ms)	TE (ms)	Resolution (mm)	Slice thickness (mm)
ADC	5194 (2486,8575)	79 (48,100)	[1.4,1.4,4.4] ([0.8,0.8,2.5],[2.2,2.2,5])	3.9 (2.5, 5)
DTI	6824 (2486,12000)	80 (48,101)	[1.9,1.9,3.5] ([0.9,0.9,2.5],[2.2,2.2,5])	3.1 (2.5, 4.2)
T2	4765 (3000,8162)	86 (77,100)	[0.5,0.5,4.3] ([0.4,0.4,3.9], [0.5,0.5,5])	3.7 (3, 4)
FLAIR	10154 (8002,11000)	113 (89,136)	[0.6,0.6,5.2] ([0.5,0.5,4.4], [1,1,6.5])	4.4 (4, 5)
Pre-T1	600 (373,819)	13 (8,18)	[0.6,0.6,4.8] ([0.5,0.5,4.4], 0.9,0.9,6.5])	4.1 (4, 5)
Post-T1	508 (250,676)	14 (2.3,21)	[0.6,0.6,4.9] ([0.5,0.5,4.4], [0.9,0.9,6.5])	4.0 (3, 5)

TI of FLAIR = 2581 (2000, 2800) ms.

7.3 Results and Discussion

After the process of data pre-processing analysis, tumour regions of conventional and diffusion MR images are defined using T2 images as references. Figure 7.1 illustrates the effectiveness of using the T2 tumour boundary across the six other image types using a typical case drawn from each of the three tumour types.

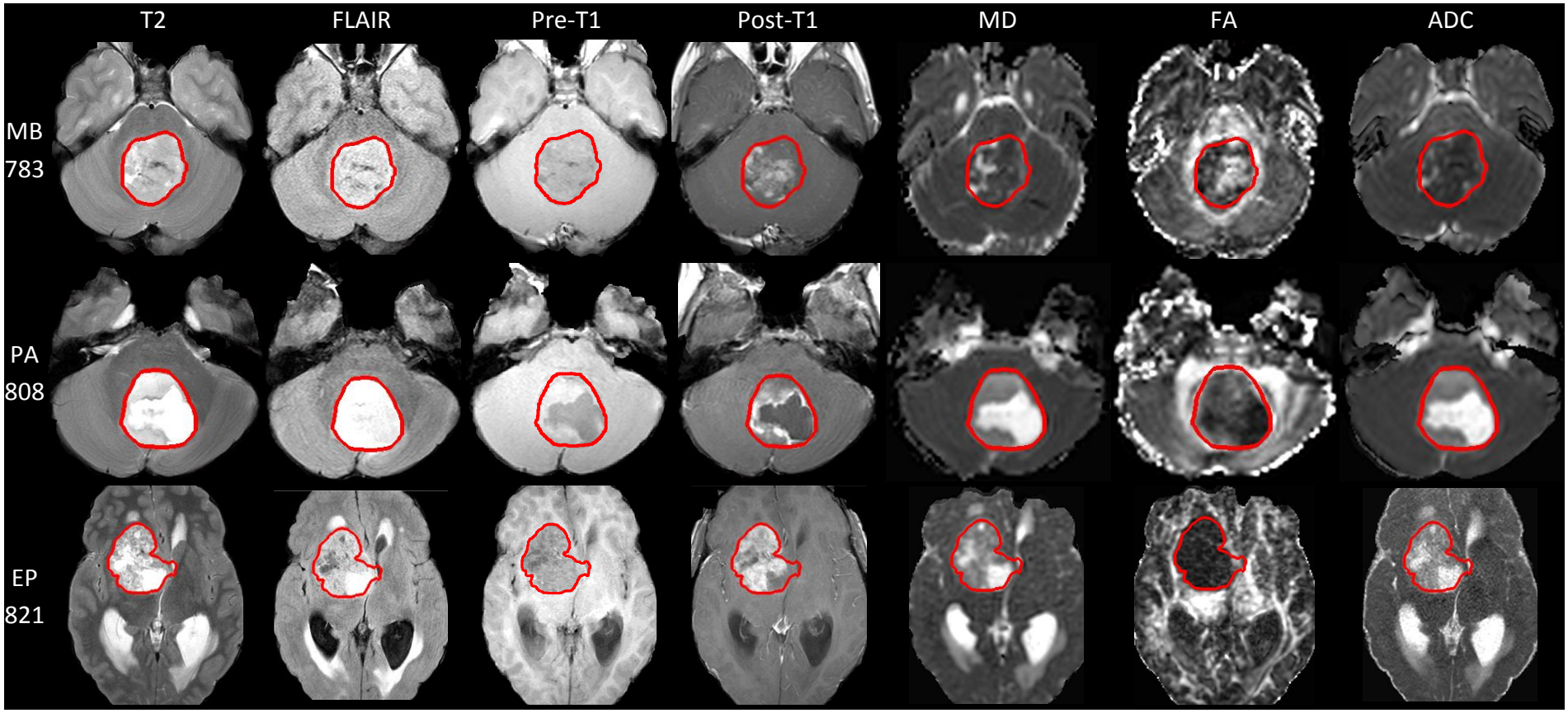


Figure 7.1: Tumour boundaries of the three tumour types. Tumour boundaries of medulloblastoma, pilocytic astrocytoma and ependymoma based on T2, FLAIR, pre-contrast T1, post-contrast T1, MD, FA, and ADC. A typical case of each tumour type is shown: cases MB783, PA808 and EP821.

Specifically, the tumour boundaries of medulloblastoma, pilocytic astrocytoma and ependymoma are drawn on seven MR-based image types: T2, FLAIR, pre-contrast T1, post-contrast T1, MD, FA and ADC. The textual features are derived from the tumour regions and subsequently an optimal set of textural features is selected for classification.

Further results and discussion in this section is organised using the following sub-sections: 1) classification based on PCA, 2) classification based on the combined mRMR and SFS method, 3) comparative analysis of the feature selection methods in order to identify which image types provide better classification outcomes, 4) evaluation of classifier, 5) effects from texture analysis and 6) comparative studies between our and the two recent published papers.

7.3.1 PCA Based Classification

From the classification based PCA method, two observations can be drawn from the results. First, conventional and diffusion MRI require different number of PCs to obtain the highest classification accuracy. The number of PCs is an important parameter in PCA analysis, with diffusion MRI generally requiring less number of components to achieve maximum classification rates than MRI (Figure 7.2). This shows that the textural features derived from diffusion MRI contain a stronger textural feature set in the top-ranked PCs.

Second, at the maximum classification accuracy, PCA based feature selection shows that ADC, T2 and post-contrast T1 based textual features provide the top-three best classification of the three tumour types, while MD, FA, pre-contrast T1 and FLAIR based textural features yield lesser accuracy in that order. The sensitivity and specificity of the

three tumour types obtained at the optimal number of PCs also shows a higher sensitivity and specificity of ADC, T2 and post-contrast T1 than other image types (Table 7.3).

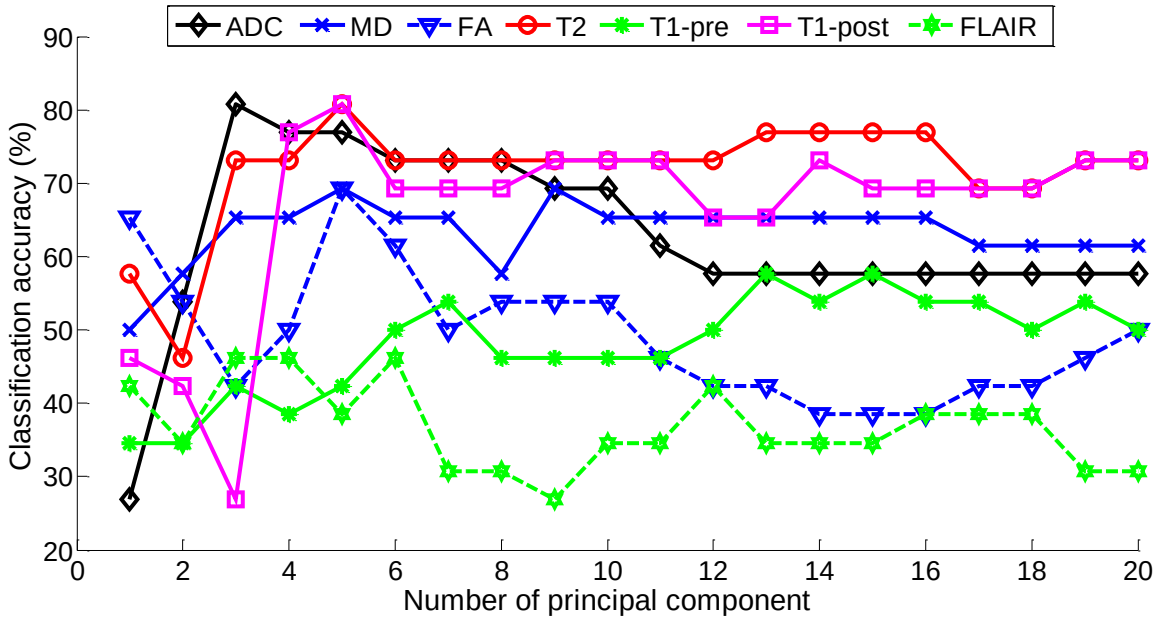


Figure 7.2: The variation of classification accuracy with respective to the number of PCs for TA based on conventional and diffusion MRI. The classification accuracy is measured using leave-one-out cross validation technique.

Table 7.3: Sensitivity and specificity for medulloblastoma, pilocytic astrocytoma, ependymoma at the optimal number of principal components.

Image	Number of PCs	Medulloblastoma		Pilocytic Astrocytoma		Ependymoma	
		Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
ADC	3	91.7%	92.9%	77.8%	94.1%	60.0%	85.7%
MD	5	72.7%	73.3%	85.7%	89.5%	50.0%	88.9%
FA	5	71.4%	83.3%	66.7%	88.2%	66.7%	82.6%
T2	5	90.1%	86.7%	87.5%	94.4%	57.1%	89.5%
FLAIR	6	46.7%	54.6%	50.0%	75.0%	20.0%	76.2%
Pre-T1	13	80.0%	75.0%	62.5%	83.3%	25.0%	77.8%
Post-T1	5	73.3%	90.9%	87.5%	94.4%	100%	87.0%

7.3.2 Combined mRMR and SFS Method Based Classification

For each image type, the ten top-ranked features are selected by mRMR method and an optimal feature set is searched by SFS technique (Table 7.4). The feature subset selected from this technique is similar to the findings presented by Krier et al. (261) in the sense that the group of most individually relevant features does not always provide the superior prediction result. Adding some less relevant features to the feature set contributes more to the accuracy of the classifier than using the top-ranked relevant features only.

This combined mRMR and SFS approach shows that the classification of ADC, T2 and post-contrast T1 based textural features are the top-three most informative, while MD, FA, FLAIR and pre-contrast T1 based textural features yield lesser accuracy in that order. This method also shows a higher sensitivity and specificity of ADC, T2 and post-contrast T1 compared with other image types (Table 7.5).

Table 7.4: Textural features selected by using mRMR and followed by SFS method.

Image type	The top 10th features selected by using mRMR	Features selected by mRMR and SFS method
ADC (DWI or DTI)	$autoc(1,0°)$, $Haar(D4)$, LRE $0°$, m, $svar(1,0°)$, LRE $45°$, $maxpr(1,0°)$, $sosv(1,0°)$, LRE $90°$, GLN $0°$	$autoc(1,0°)$, GLN $0°$, $sosv(1,0°)$
MD (DTI)	$savg(5,0°)$, $dent(2,45°)$, $svar(1,0°)$, $SRLGE$ $90°$, m, SRE $45°$, $svar(1,45°)$, $maxpr(3,135°)$, SRE $135°$, $svar(1,90°)$	$svar(1,0°)$, $svar(1,45°)$, $svar(1,90°)$, m, SRE $135°$, $SRLGE$ $90°$, $maxpr(3,135°)$, $savg(5,0°)$
FA (DTI)	$entro(2,45°)$, $cshad(5,135°)$, $sym4(D4)$, $corrp(5,45°)$, $sym4(H3)$, RP $90°$, $entro(2,90°)$, $corrp(3,0°)$, $SRHGE$ $45°$, $entro(3,45°)$	RP $90°$, $SRHGE$ $45°$
T2	$cshad(3,135°)$, $sent(5,135°)$, $LRLGE$ $90°$, $cshad(1,90°)$, m, $HGRE$ $0°$, $cshad(2,135°)$, $maxpr(2,135°)$, $cshad(3,0°)$, $LRHGE$ $0°$	$cshad(3,135°)$, $sent(5,135°)$, m, $HGRE$ $0°$, $cshad(2,135°)$, $cshad(3,0°)$
FLAIR	$dent(3,90°)$, $sym4(H1)$, GLN $135°$, $imc1(2,90°)$, $dent(4,0°)$, $D4(D2)$, GLN $90°$, $imc1(3,90°)$, $homo(5,90°)$, var	$sym4(H1)$, $D4(D2)$, GLN $135°$, $imc1(3,90°)$, GLN $90°$, var
Pre-contrast T1	$sym4(H1)$, $sent(5,90°)$, $LGRE$ $0°$, var, sk, $Haar(H2)$, $imc2(1,90°)$, $entro(1,0°)$, $Haar(A)$, $imc1(1,135°)$	$sym4(H1)$
Post-contrast T1	$sym4(D1)$, $cprom(1,45°)$, GLN $0°$, $cshad(2,45°)$, $sym4(H1)$, $LRLGE$ $0°$, $SRLGE$ $45°$, $cshad$ $(3,90°)$, GLN $45°$, $cprom(1,135°)$	$sym4(D1)$, GLN $0°$, $cshad$ $(3,90°)$,$cshad(2,45°)$, $cprom(1,45°)$

Table 7.5: Sensitivity and specificity for medulloblastoma, pilocytic astrocytoma, ependymoma derived from the combined mRMR and SFS method based classification.

Image type	Medulloblastoma		Pilocytic Astrocytoma		Ependymoma	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
ADC	100%	92.9%	100%	94.4%	66.7%	100%
MD	91.7%	85.7%	100%	94.4%	50.0%	95.0%
FA	83.3%	85.7%	87.5%	77.8%	33.3%	95.0%
T2	91.7%	92.9%	100%	94.4%	83.3%	100%
FLAIR	100%	35.7%	50.0%	100%	16.7%	100%
Pre-T1	100%	28.6%	50.0%	100%	0%	100%
Post-T1	100%	92.9%	100%	88.9%	50.0%	100%

In order to visualise higher dimensional variables, we employed PCA to study the distribution of features in two-dimensional space. After features are selected by the combined mRMR and SFS method, the features are transformed to the 1st and 2nd PC (see Figure 7.3). This figure illustrates that medulloblastoma and pilocytic astrocytoma can be linearly separated in the cases of ADC, MD, T2 and post-contrast T1 images but less likely to be linearly separable for other image types.

Features obtained from ependymoma images lie between those obtained from medulloblastoma and pilocytic astrocytoma images and therefore difficult to differentiate from the other two tumour types. It is important to note here that although this figure provides a useful graphical explanation of the classification difficulty using 2D projection, the classification results presented in the thesis are based on a higher-dimensional classification exercise.

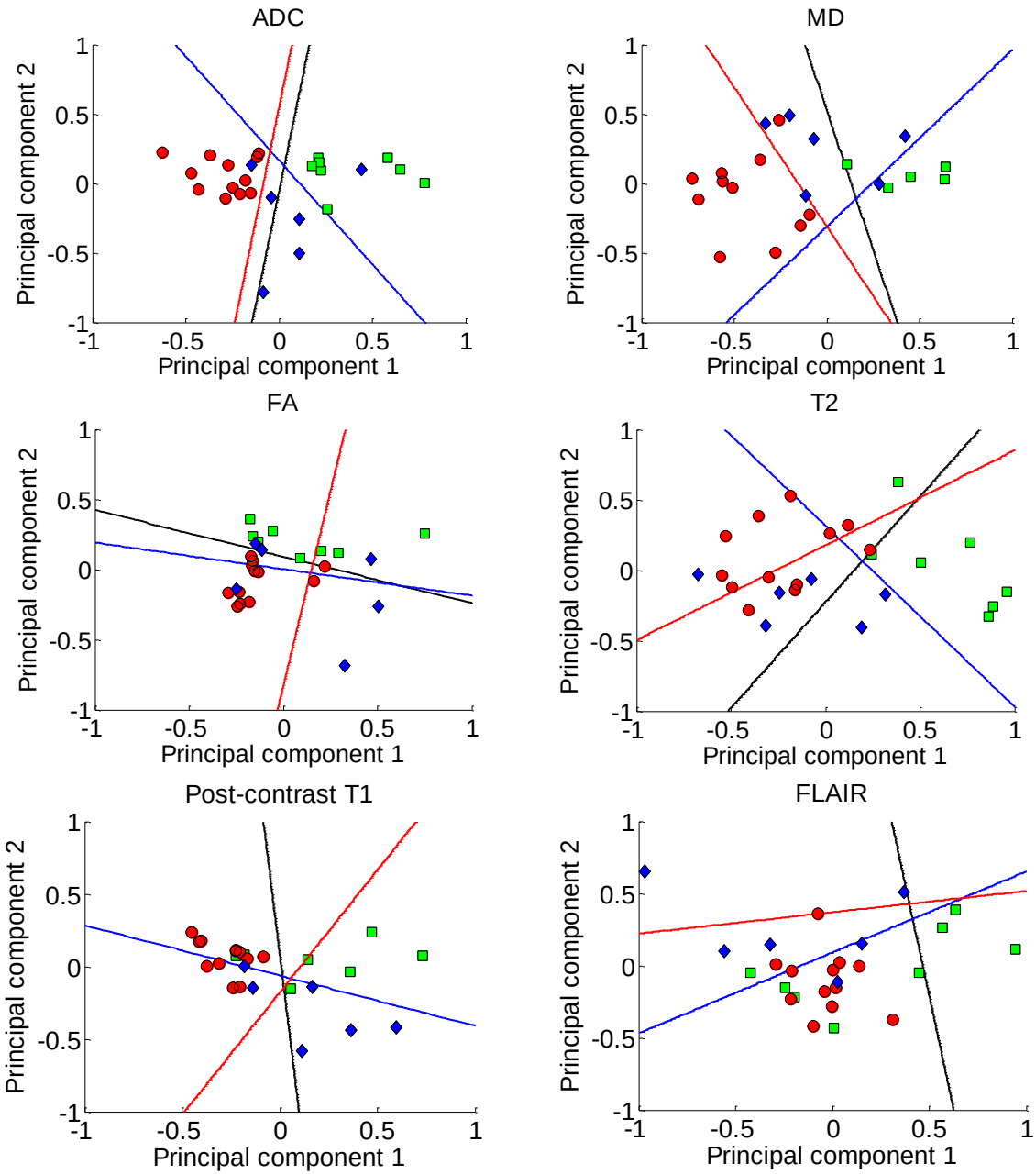


Figure 7.3: Plot of 1st and 2nd PC of features selected by combined mRMR and SFS method.

Legend: MB PA EP

Hyperplane separating MB and PA

Hyperplane separating MB and EP

Hyperplane separating PA and EP

7.3.3 Comparative Analysis

In terms of classification accuracy, PCA based classification generally provides better results than using the original feature set, but gives lower discrimination rates than the combined mRMR and SFS method (Figure 7.4). The fact is that the PCA method transforms all training data to obtain the maximum variance and still keeps all textural feature variables to generate the PCs. Some outliers or noise, contributed by less separable textural features in the PCs, can affect the classification outcome. Combining mRMR and SFS schemes enhances the discrimination performance for all types of MRI.

In statistical hypothesis testing (Table 7.6), PCA and the combined mRMR and SFS method based classification are not significantly different at 90% confidence interval, except the combined mRMR and SFS method produces significantly higher classification accuracy for FLAIR at 90% confidence interval.

However, the combined mRMR and SFS method based classification produce relatively better result than PCA based classification when compared with original features based classification. The combined mRMR and SFS method based classification produces significantly higher classification results than the original features based classification for ADC, FA, FLAIR and post-contrast T1 at 95% confidence interval and for MD and T2 at 90% confidence interval. The PCA based classification can only produces better classification for ADC and FA at 90% confidence interval.

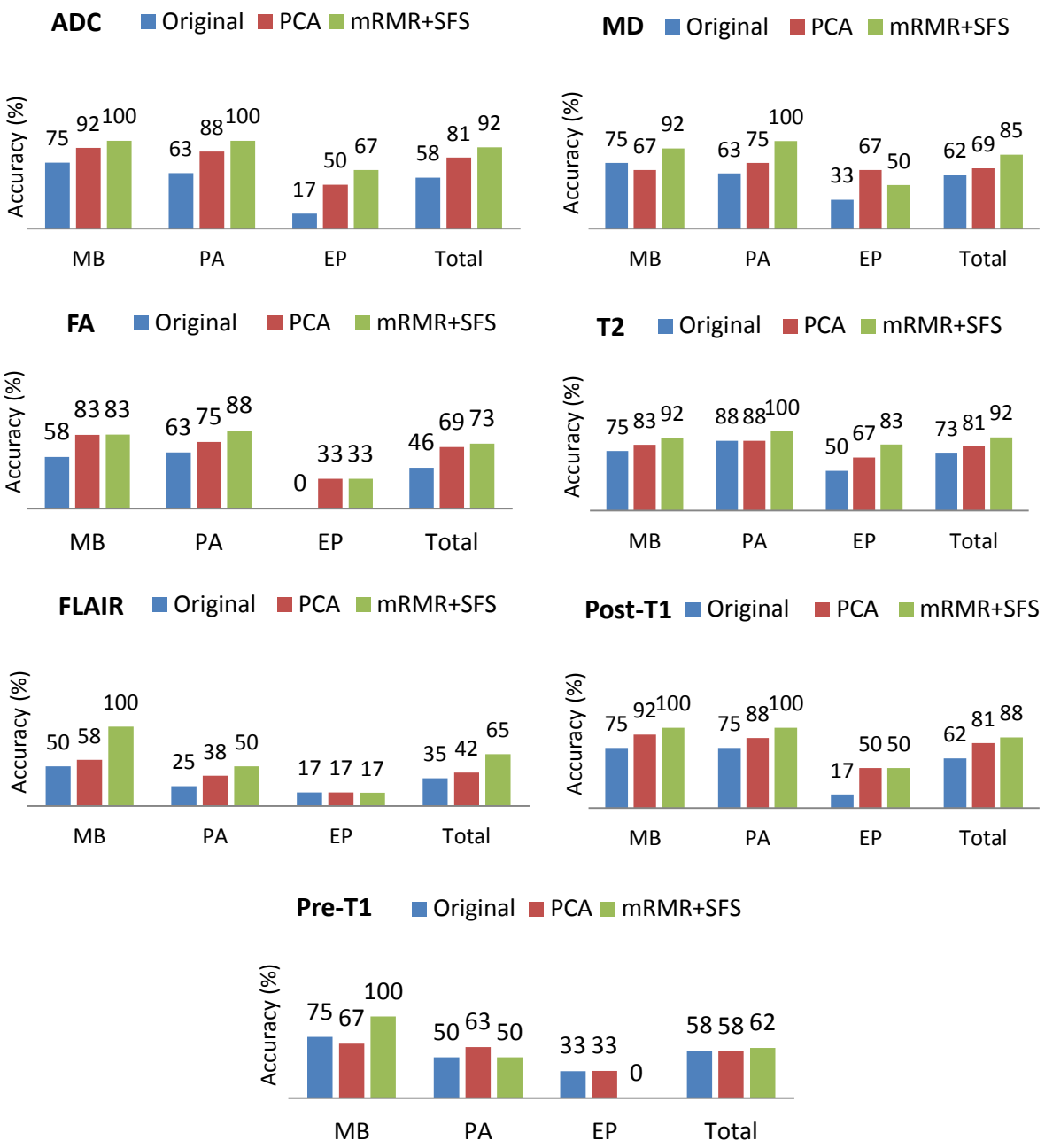


Figure 7.4: Classification accuracy of MB, PA, EP and overall tumour types obtained from PCA, combined mRMR and SFS and original features based classification The classification accuracy values are obtained for ADC, MD, FA, T2, FLAIR, pre-contrast T1 and post-contrast T1, where the accuracy figures are measured using LOOCV technique.

Table 7.6: P-values of statistical hypothesis test between the classification accuracy obtained from three pairwise comparisons. 1) Original features vs PCA based classification. 2) Original features vs the combined mRMR and SFS method based classification. 3) PCA and the combined mRMR and SFS method based classification. The bold figures are p-values showing significant results at either 95% confidence interval (p-value < 0.05) or 90% confidence interval (p-value < 0.1).

Image types	ADC	MD	FA	T2	FLAIR	Pre-T1	Post-T1
Original vs PCA	**0.0714**	0.5599	**0.0921**	0.5103	0.5686	1	0.1259
Original vs mRMR+SFS	**0.0039**	**0.0607**	**0.0479**	**0.0668**	**0.0265**	0.7775	**0.025**
PCA vs mRMR+SFS	0.2229	0.188	0.7595	0.2229	**0.0951**	0.7775	0.4421

Despite the statistically insignificant difference of the two methods, the combined mRMR and SFS method based classification is observed to improve the discriminative performance of childhood brain tumours for all image types. With the combined mRMR and SFS method, the classification is examined using four validation methods. The LOOCV $\widehat{Err}^{(cv1)}$, the LOOB $\widehat{Err}^{(1)}$, the .632 estimator $\widehat{Err}^{(.632)}$ and the .632+ estimator $\widehat{Err}^{(.632+)}$ are computed to estimate classification error of specific MR-based image type, as shown in Table 7.7. The .632+ estimator reduces bias and gives the error estimation values in the range bounded by the error values obtained from LOOCV, LOOB and .632 estimators. The four validation methods show that T2, ADC, post-contrast T1 provide better outcome than the other image types, matching the results obtained from PCA based classification.

Table 7.7: Comparison of classification error estimated from LOOCV, LOOB, .632 estimator and .632+ estimator. The error values range between zero and one; a higher error value refers to lesser power of discrimination.

Image type	Validation Techniques			
	$\widehat{Err}^{(cv1)}$	$\widehat{Err}^{(1)}$	$\widehat{Err}^{(.632)}$	$\widehat{Err}^{(.632+)}$
ADC	0.0769	0.0892	0.0847	0.0848
MD	0.1538	0.1608	0.1158	0.1220
FA	0.2692	0.2769	0.2599	0.2612
T2	0.0769	0.0669	0.0564	0.0568
FLAIR	0.3462	0.3492	0.3481	0.3481
Pre-T1	0.3846	0.3646	0.3720	0.3723
Post-T1	0.1154	0.1146	0.1149	0.1149

In order to identify which image type can provide better classification outcome, the p-values of comparative statistical testing are measured and compared between classification accuracy obtained from different MR-based image types, as shown in Table 7.8. Four observations can be drawn from these p-values. First, at 95% confidence interval, the classification accuracy obtained from ADC, T2 and post-contrast T1 based are significantly better than those obtained from FLAIR and pre-contrast T1 based TA.

Second, at 90% confidence interval, the classification accuracy obtained from ADC and T2 based TA is significantly higher than that from FA based TA. However, the classification accuracy obtained from post-contrast T1 based TA is not significantly better than that from FA based TA. Third, at 95% confidence interval, the classification accuracy derived from ADC, T2 and post-contrast T1 do not show significantly different results among themselves.

Fourth, the classification accuracy obtained from MD based TA does not show significantly different result compared to ADC, T2 and post-contrast T1 based TA. However, their p-values identify that MD based TA tends to provide less power of discrimination among the tumour types than ADC, T2 and post-contrast T1 based TA. As a result, ADC, T2 and post-contrast T1 based TA are selected for analysing multimodal MRI based TA, as demonstrated in chapter 8.

Table 7.8: P-values of statistical hypothesis test between the classification accuracy obtained from two MR-based image types. The bold figures are p-values showing significant results at either 95% confidence interval (p-value < 0.05) or 90% confidence interval (p-value < 0.1).

	ADC	MD	FA	T2	FLAIR	Pre-T1	Post-T1
ADC	1	0.3853	**0.0668**	1	**0.0175**	**0.0085**	0.6381
MD	0.3853	1	0.3083	0.3853	0.1093	**0.0607**	0.6845
FA	**0.0668**	0.3083	1	**0.0668**	0.5479	0.3751	0.1593
T2	1	0.3853	**0.0668**	1	**0.0175**	**0.0085**	0.6381
FLAIR	**0.0175**	0.1093	0.5479	**0.0175**	1	0.7733	**0.0483**
Pre-T1	**0.0085**	**0.0607**	0.3751	**0.0085**	0.7733	1	**0.025**
Post-T1	0.6381	0.6845	0.1593	0.6381	**0.0483**	**0.025**	1

7.3.4 Evaluation of Classifier

The risk of over-fitting as mentioned in section 3.9 is addressed by using two approaches. The first approach is dimensionality reduction and the second approach is bootstrapping and cross-validation methods.

During the feature selection phase, dimensionality reduction techniques are employed such that the number of chosen features is kept small relative to the number of samples. It can be observed that the numbers of features selected using the mRMR and SFS method are 3 for ADC, 6 for T2 and 5 for post-contrast T1. In addition, the numbers of PCs that provide highest average classification accuracy are 3 for ADC and 5 for both T2 and post-contrast T1.

The use of bootstrapping and cross-validation statistics suggest that the three-class classifier of ADC, T2 and post-contrast T1 based TA have a high predictive ability for future cases of these tumour types. The classification accuracy rates obtained from four validation techniques are 91%-92% for ADC, 92%-94% for T2 and 88%-89% for post-contrast T1.

7.3.5 Effects from Texture Analysis

The number of grey levels used in GLCM and GLRLM has an impact on the texture patterns and can affect the classification outcome. This is because the number of grey levels can alter the elements of matrices, hence the textural features that depend on them. Inappropriate quantization, i.e. too low and too high number of grey levels, can provide unsatisfactory results. Furthermore, a large number of grey levels requires a longer processing time. The following section discusses the effect of the number of grey levels to the classification accuracy derived from GLCM and GLRLM. In addition, two other effects from using texture analysis: shift-variance and three-dimensional texture analysis are also discussed.

In our literature review, there was only one direct reference to the number of grey levels employed for TA of MRI based brain tumour classification. In Kassner et al. (18), the number of grey levels was reported to be scaled down to either 32 or 64. Experimentally, using 64 grey-levels requires extensive computation time on a standard PC. As result, we calculated the classification accuracy against number of grey levels ranging from 4 to 32 used in GLCM and GLRLM, as shown in Figure 7.5 and 7.6.

The observed classification accuracy does not vary linearly with the number of grey levels used for calculating the textural features. In fact, the pattern of variation is highly irregular. The optimal number of grey levels for GLCM and GLRLM (Table 7.9) do not exceed 32 and are less than 16 in most cases, except for FLAIR based TA. At the optimal number of grey levels from Table 7.9, the classification accuracy is calculated with the use of combined textural features obtained from first-order statistics, absolute gradient, GLCM, GLRLM, Haar, Daubechies 4 and Symlet 4.

In Table 7.10, the classification accuracy obtained by using the optimal number of grey levels is similar compared to the classification accuracy presented in section 7.3.3, which employs nine grey-levels. However with the optimal number of grey levels, a small gain for MD and larger increase for pre-contrast T1 are observed.

As the sample size used in this section is small, the optimal number of grey levels obtained in experiment 3 may not be applicable for a different dataset. In addition, altering the number of grey levels may not affect the classification outcome when combining textural features from different texture analysis approaches is used, as observed from TA of some image types in Table 7.10.

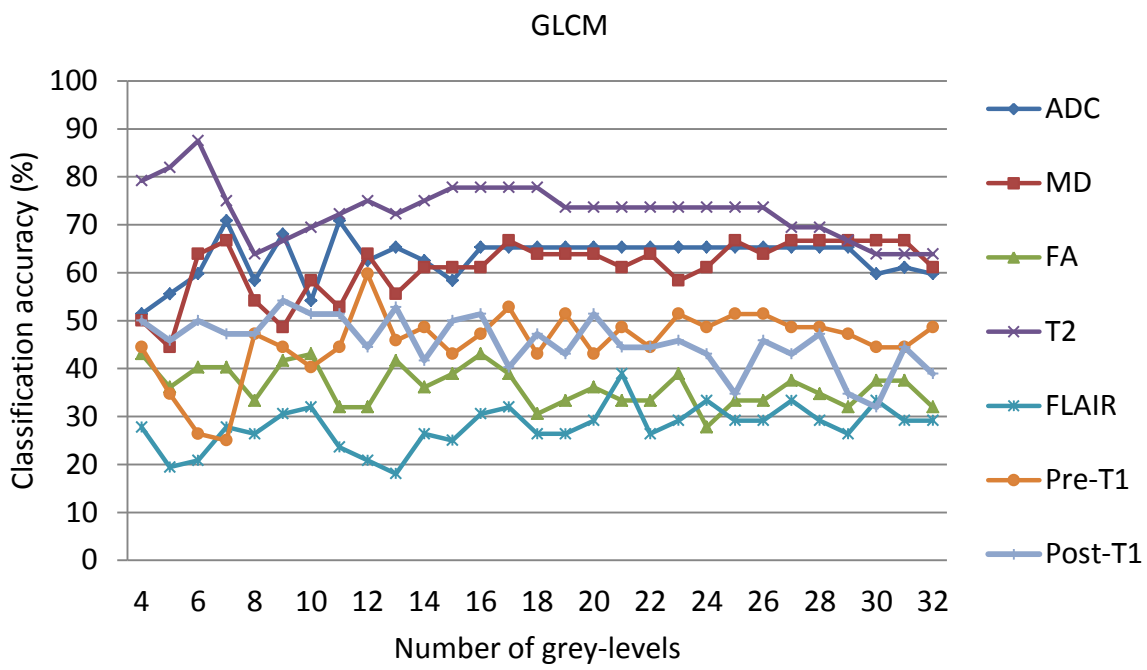


Figure 7.5: The variation of classification accuracy with respect to the number of grey levels for GLCM based TA.

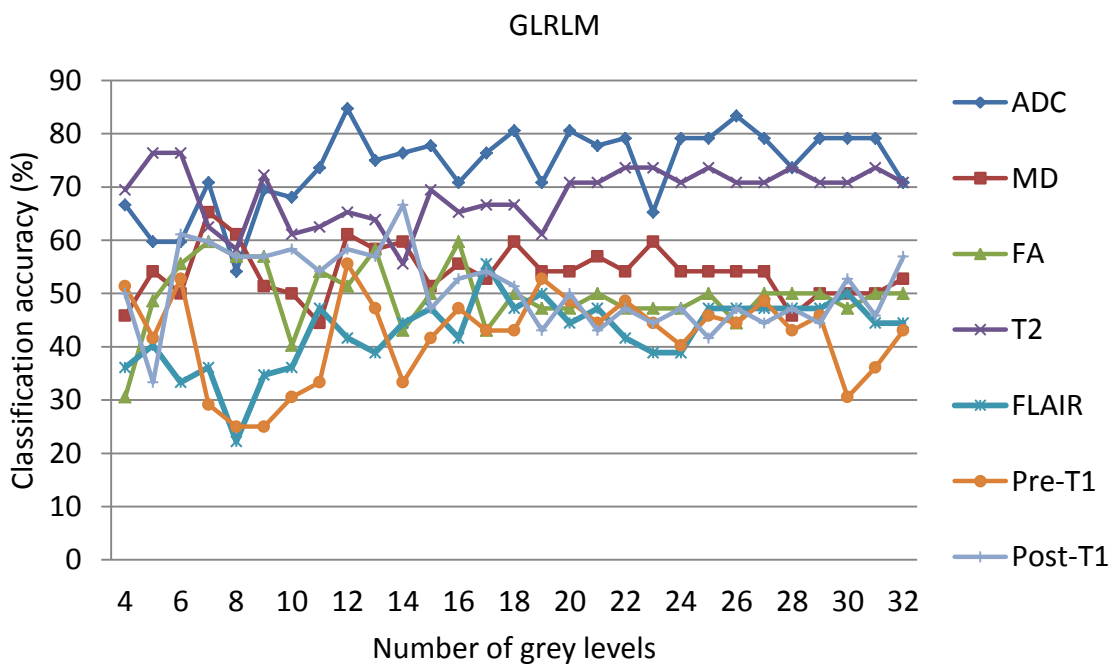


Figure 7.6: The variation of classification accuracy with respect to the number of grey levels for GLRLM based TA.

Table 7.9: The optimal number of grey levels that gives the highest classification accuracy derived from GLCM and GLRLM.

Method	ADC	MD	FA	T2	FLAIR	Pre-T1	Post-T1
GLCM	7	7	4	6	21	12	9
GLRLM	12	7	7	5	17	12	14

Table 7.10: Classification accuracy obtained from the optimal number of grey levels vs nine grey levels. The classification is based on textural features obtained from first-order statistics, absolute gradient matrix, GLCM, GLRLM, Haar, Daubechies 4 and Symlet 4.

Image type	ADC	MD	FA	T2	FLAIR	Pre-T1	Post-T1
Optimal grey-levels	92%	88%	73%	92%	65%	85%	88%
Nine grey-levels	92%	85%	73%	92%	65%	62%	88%

Apart from variation of the number of grey levels in statistical texture analysis, shift-invariant texture and higher dimensional texture analysis should also be mentioned. Texture matrices which are altered by rotating, translating and scaling of an image, can also affect the classification outcome. For example, DWT is shift-variant and its texture matrices extracted from wavelet coefficients will also be shift-variant, and therefore can degrade the classification performance (295-297).

Although the two-dimensional TA has shown promising results in classifying the primary childhood brain tumours, work on three-dimensional TA has revealed promising results for the discrimination of secondary brain tumours (298). Three-dimensional TA also offers better discrimination between necrosis and solid tumours as well as between oedema and solid tumours (180).

Consequently, the effects of shift-invariant texture matrices and three-dimensional TA on discrimination of the primary brain tumours should be inspected for future work.

7.3.6 Comparative Studies

In this section, we discuss the results obtained from our findings and two recent published papers from Vlachou et al. (19) and Rodriguez Gutierrez et al. (30). It should be noted that these studies are not based on the same study cases; therefore the figures listed in following tables are not directly comparable.

Vlachou et al. (19) presented the classification of posterior fossa tumours based TA. The sensitivity and specificity obtained from T2 and pre-contrast T1 based TA are shown in Table 7.11. Our T2 based TA result produced sensitivity and specificity for the three tumour types as high as the PNN based classification and better than the LDA based classification in Vlachou et al. (19). However our pre-contrast T1 based TA tends to produce lower classification accuracy compared to PNN based classification, but show similar outcomes to LDA based classification in Vlachou et al. (19).

Table 7.11: Sensitivity and specificity derived from T2 and pre-contrast T1 based TA, reported in Vlachou et al. (19) and our study.

	T2			Pre-contrast T1		
	MB	PA	EP	MB	PA	EP
PNN based classification presented in Vlachou et al. (19)						
Sensitivity (%)	95.2	95.2	80.0	96.8	90.5	60.0
Specificity (%)	91.2	98.7	98.1	84.2	96.2	100
LDA based classification presented in Vlachou et al. (19)						
Sensitivity (%)	58.7	71.4	13.3	36.5	50.0	6.7
Specificity (%)	82.5	87.2	70.5	63.2	70.5	70.5
Our study						
Sensitivity (%)	91.7	100	83.3	100	50.0	0
Specificity (%)	92.9	94.4	100	28.6	100	100

Rodriguez Gutierrez et al. (30) presented the classification of posterior fossa tumours based on histogram analysis of ADC values. The classificiation accuracy derived from their and our studies is shown in Table 7.12. The classification of ependymoma in their study tends to provide higher accuracy.

The result differences between ours and the two recent publications can be affected by several factors, which are discussed in section 3.11. However, it is beyond the scope of this thesis to verify whether such differences are attributable to the dataset used or the use of different methods. It should also be noted that we present a semi-automatic processing pipeline which aims to improve diagnostic classification of brain tumours holistically rather than focussing solely on classification accuracy. Furthermore, the framework is flexible to permit the integration of different techniques or optimisations.

Table 7.12: Comparison of classification accuracy derived from ADC based TA, reported in Rodriguez Gutierrez et al. (30) and our study.

	MB	PA	EP
Rodriguez Gutierrez et al. (30)			
Classification accuracy (%)	95.8	96.9	94.3
Our study			
Classification accuracy (%)	100	100	67

7.4 Conclusions

This chapter presented the examination of single MRI modality based TA and identified which MR-based image type can better discriminate childhood brain tumour types. Three conclusions were drawn from experiment 3.

First, amongst all data from various types of conventional and diffusion MRI obtained from the multiple centres, the combined mRMR and SFS method provides a textural feature subset that are observed to better discriminate the three tumour types than using the original features or the transformed features extracted by using the PCA.

Second, with the combined mRMR and SFS method, classification based on textural features of ADC, T2-weighted and post-contrast T1-weighted images provide an overall classification accuracy of 92%, 92% and 88% respectively and significantly outperform textural features of FLAIR and pre-contrast T1. They are also observed to give higher power of discrimination than DTI based TA.

Third, the ADC, T2 and post-contrast T1 based TA indicate a good separation between medulloblastoma and pilocytic astrocytoma. However, an improvement in the discrimination of ependymoma from the other two groups is required and suggested as a future study.

In the next chapter, combining textural features based on multimodal MRI is investigated with larger samples of the multi-centre dataset.

CHAPTER 8 EXPERIMENT 4: CLASSIFICATION OF BRAIN TUMOURS BASED ON TEXTURE ANALYSIS OF MULTIMODAL MRI

To meet the final objective of section 1.2, this chapter demonstrates the combination of conventional and diffusion MRI based TA of a multi-centre dataset. The classification of childhood brain tumours based on TA of single modality and multimodal MRI are compared. Experiment 4 follows the flowchart in section 4.5. The automated processing pipeline described in chapter 5 to 7 is adapted with the use of a multinomial logistic regression method and the LOOCV technique as the examination criteria.

8.1 Introduction

As reviewed in section 3.4.2, single modality MR image analysis is not sufficient to completely represent histologic types of brain tumours. The use of multimodal MRI has been shown to provide more comprehensive characterisation of brain tumours. Three common types of MRI: ADC, T2-weighted and post-contrast T1-weighted images have been shown to exhibit high sensitivity and specificity in differentiating brain tumours in several studies (30, 196-200, 299) and in chapter 7 of this thesis. These types of MRI have their own unique advantages. ADC measures the magnitude of water diffusion, which helps to characterise the posterior fossa tumours with high sensitivity. T2 provides an indication of free water abundance and macromolecules, which offer high sensitivity of

distinction between medulloblastomas and pilocytic astrocytomas. Post-contrast T1 enhances pathology and exhibits the region with blood-brain barrier[11] breakdown.

As discussed above, it is hypothesised that integrating the advantages of ADC, T2-weighted and post-contrast T1-weighted images, through the use of TA based on statistical and wavelet transform approaches, would offer more comprehensive discrimination of childhood brain tumours. In order to examine this hypothesis, the automated processing pipeline follows the processes as aforementioned in chapter 5 to 7. However, a classifier that can provide probability values, is used instead of the SVM method to examine multimodal MRI based TA. A multinomial logistic regression is one of popular probability classification algorithms and is commonly used for medical data analysis (181, 191). This method is thus employed as a classifier and validated by the LOOCV technique.

8.2 Material

A pre-treatment MRI dataset of paediatric brain tumours was considered. The number of embryonal, astrocytic and ependymal cases available at the time this study was undertaken were 23, 29, and 19 cases and their subtypes are shown in Table 8.1. The acquisition of T2-weighted, post-contrast T1-weighted and DW images are acquired from five recruitment centres using 1.5T and 3T Siemens, 1.5T GE, 1.5T and 3T Philips scanners. The average and range of TR, TE, TI, resolution and slice thickness of the brain imaging data are shown in Table 8.2. The study cases with value of TR, TE, TI, resolution, FOV, slice

[11]Blood-brain barrier is the permeability of brain capillaries controlling the type and amount of substances passing to the blood vessel.

thickness and magnetic field strength as well as scanner's manufacture used in experiment 4 are listed in section C.4, Appendix C.

Table 8.1: Number of subjects for each tumour type for experiment 4.

WHO Defined Tumour	Embryonal		Astrocytic	Ependymal	
Subtype	medulloblastoma	Desmoplastic /nodular medulloblastoma	Pilocytic Astrocytoma	Ependymoma	Anaplastic Ependymoma
Number of cases	21	2	29	6	13

Table 8.2: Imaging parameters for experiment 4. Average and range of TR, TE, resolution and slice thickness are presented in the format of average (min,max).

Image type	TR (ms)	TE (ms)	Resolution (mm)	Slice Thickness (mm)
ADC	5175 (2486, 8900)	89 (48, 161)	[1.4, 1.4, 5.2] ([0.8, 0.8, 2.5], [2.4, 2.4, 8])	4.3 (2.5, 5)
T2	4916 (2660, 9140)	88 (14, 119)	[0.5, 0.5, 5.1] ([0.4, 0.4, 3.9], [0.9, 0.9, 8.5])	4.0 (3, 5)
Post-T1	587 (8, 2640)	15 (2.3, 23)	[0.6, 1.6, 5.4] ([0.4, 0.4, 1.9], [0.9, 0.9, 7.9])	4.3 (0.8, 6)

8.3 Methods

From the results of chapter 7, the classification of ADC, T2 and post-contrast T1 based TA are conducted in order to compare with the classification of multimodal MRI based TA. In experiment 4, multinomial logistic regression is used as a classifier, as reviewed in section 3.7. The deviance values derived from the models are reported to relatively compare the quality of the fitted models. In addition, classification accuracy obtained from single

modality and multimodal MRI based TA approaches are statistically compared using a hypothesis test of two proportions. Three approaches of classification of multimodal MRI based textural features are defined. The first approach selects the first ten textural features ranked by using the mRMR method for ADC, T2 and post-contrast T1 and concatenates these textural features into a feature vector. After an optimal feature set is obtained from the concatenated feature set using the SFS method, the class of data is predicted. This approach is named as the combined multimodal MRI based TA, as shown in Figure 8.1.

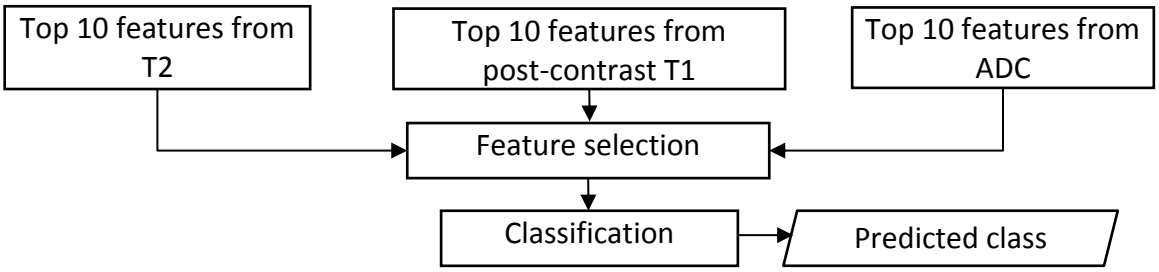


Figure 8.1: Classification based on combined multimodal MRI based TA.

The second approach combines the probability of predicted classes derived from each image type, as shown in Figure 8.2.

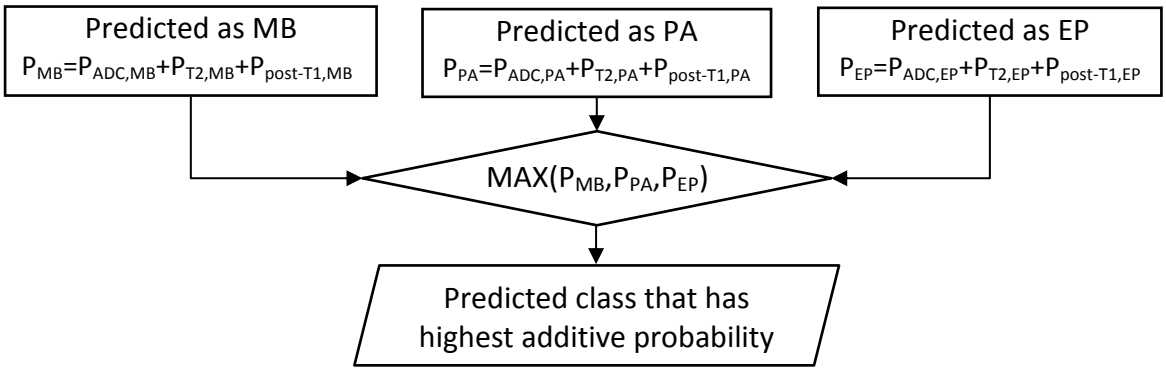


Figure 8.2: Classification based on additive probability outcome.

For example, the probability of being medulloblastoma obtained from individual ADC, T2 and post-contrast T1 are denoted as $P_{ADC,MB}$, $P_{T2,MB}$ and $P_{post\text{-}T1,MB}$ respectively. Subsequently, the probability of being medulloblastoma derived from these image types are $P_{MB} = P_{ADC,MB} + P_{T2,MB} + P_{post\text{-}T1,MB.}$ Consequently, the highest additive probability determines the final predicted class. This approach is named as the additive probability outcome.

The third approach examines the predicted class for each image type of T2, post-contrast T1 and ADC and assigns a weight (overall accuracy of each image type) to each predicted class. The outcome of this approach is a class having the highest vote from the weighted combination of predicted classes. This approach is named as the superposition of weighted outcomes (Figure 8.3).

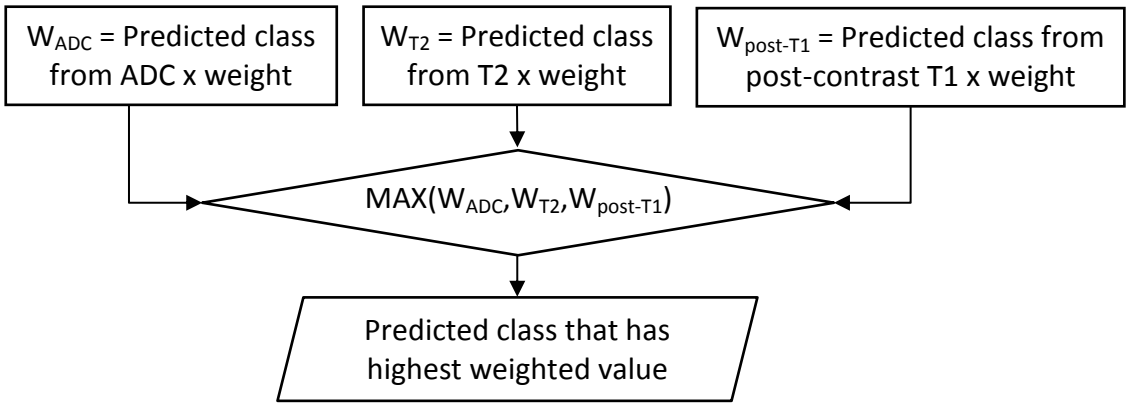


Figure 8.3: Classification based on superposition of weighted outcomes.

8.4 Results and Discussion

This section is organised using the following sub-sections: 1) examining textural features selected from ADC, T2, post-contrast T1 and multimodal MR images, 2) classification outcomes derived from single modality and multimodal MRI based TA, 3) statistical

analysis comparing different classification approaches, 4) visual examination of misclassified cases, 5) building multinomial logistic regression models of suggested processing frameworks, 6) recommended processing frameworks, 7) biological perspective and 8) classification difficulties and future work.

8.4.1 Examining Textural Features

After pre-processing and tumour segmentation described in chapter 5 and 6, high dimensionality of textural features is generated. In order to obtain the maximum relevant and minimum redundant textual features, the ten top-ranked features are selected by using mRMR method. In order to measure performance of these textural features for discriminating brain tumours, classification accuracy and deviance are computed, as shown in Table 8.3 - 8.5.

In general, features with lower deviance have higher classification accuracy, and vice versa. However, some features that are less relevant to the class or fit to the model less well (higher deviance) can produce higher classification accuracy. For example, in ADC based TA, $svar(2, 0°)$**,** having a deviance of 101.6, produces higher accuracy than m**,** having a deviance of 89.5 (Table 8.3). This means that the correlation between a features and actual class does not always explain the accuracy or error associated with the classification model.

In T2 and post-contrast T1 based TA (Table 8.4 and Table 8.5), their lower classification accuracy and higher deviance imply that their textural features fit the model less well, compared to the textural features of ADC. ADC based TA can distinguish all tumour types

better than T2 and post-contrast T1 based TA. However single modal features derived from none of these image types were able to achieve a satisfactory accuracy level for ependymoma cases.

Table 8.3: Classification accuracy and deviance of the ten top-ranked textural features selected by using mRMR method from ADC dataset.

ADC	Deviance	Predicted Correct (%)			
		MB	PA	EP	Overall
$savg(5, 135°)$	93.6	87.0	86.2	26.3	70.4
$idmn(2, 90°)$	143.7	0.0	69.0	42.1	39.4
$sosv(4, 0°)$	96.3	87.0	82.8	26.3	69.0
$RP(45°)$	148.5	0.0	86.2	21.1	40.8
$svar(2, 0°)$	101.6	82.6	86.2	36.8	71.8
m	89.5	91.3	79.3	21.1	67.6
$cshad(1, 0°)$	118.9	56.5	79.3	5.3	52.1
$SRLGE(135°)$	145.1	47.8	72.4	0.0	45.1
$savg(5, 45°)$	93.7	87.0	86.2	21.1	69.0
$svar(2, 135°)$	101.9	82.6	86.2	36.8	71.8

Table 8.4: Classification accuracy and deviance of the ten top-ranked textural features selected by using mRMR method from T2 dataset.

T2	Deviance	Predicted Correct (%)			
		MB	PA	EP	Overall
$cshad(5, 0°)$	113.7	78.3	75.9	10.5	59.2
$entro(1, 90°)$	141.1	4.3	75.9	47.4	45.1
$cshad(1, 90°)$	107.9	78.3	82.8	15.8	63.4
$svar(1, 90°)$	109.9	73.9	82.8	10.5	60.6
$imc1(4, 0°)$	131.4	69.6	65.5	0.0	49.3
$cshad(5, 45°)$	116.1	78.3	79.3	15.8	62.0
$dent(2, 135°)$	128.0	39.1	82.8	0.0	46.5
$savg(1, 0°)$	110.7	73.9	82.8	21.1	63.4
$cshad(2, 90°)$	109.5	78.3	82.8	15.8	63.4
$Haar(V2)$	147.1	0.0	79.3	26.3	39.4

Table 8.5: Classification accuracy and deviance of the ten top-ranked textural features selected by using mRMR method from post-contrast T1 dataset.

Post-contrast T1	Deviance	Predicted Correct (%)			
		MB	PA	EP	Overall
$cshad(2, 0°)$	143.1	65.2	65.5	0.0	47.9
$LGRE\ 0°$	139.5	43.5	48.3	26.3	40.8
var	129.9	69.6	72.4	0.0	52.1
$imc1(2, 135°)$	138.8	60.9	75.9	0.0	50.7
$SRLGE(135°)$	143.1	39.1	62.1	26.3	45.1
$Haar(V2)$	145.5	21.7	41.4	5.3	25.4
$cshad(2, 135°)$	142.9	69.6	69.6	0.0	50.7
$cprom(1, 0°)$	121.3	82.6	79.3	0.0	59.2
$RP(90°)$	147.0	0.0	86.2	42.1	46.5
$homo(1, 135°)$	146.5	0.0	86.2	21.1	40.8

After the ten top-ranked features are selected, the optimal textural feature set for ADC, T2, post-contrast T1 and the combined textural features was obtained by using SFS method, as shown in Table 8.6. The textural features selected in this study are different from the features selected in Table 7.4, section 7.3 because of the use of larger dataset containing more complicated cases and the use of the different classifier providing different discriminative criteria. (There is more number of cases that have the scans of ADC, T2 and post-contrast T1 than cases having all seven MR-based image types, used in chapter 7.)

However, there are some common textural features selected from both classifiers (SVM and MLR method). The common textural features are derived from grey-level co-occurrence matrices: $sosv$ for ADC, $cshad$ for T2-weighted and $cprom$ for post-contrast T1-weighted images. In addition, the optimal feature set contains GLCM based textural features as majority members for single modality and multimodal MRI. From this

observation, the GLCM based textural features can be considered as important features for the classification of childhood brain tumours.

In a related study, Rodriguez Gutierrez et al. (30) reported that in their ADC based TA study, sum average and sum variance based GLCM are the most interesting textural features apart from texture based histogram. Our study also found that these two textural features ($svar$, $savg$) are the top most discriminative features for ADC based TA, as shown in Table 8.6.

From Table 8.6, it is observed that the optimal numbers of features are 6 for both ADC and T2, 4 for post-contrast T1 and 10 for the combined ADC, T2, and post-contrast T1. The number of features chosen based on the mRMR and SFS method is smaller than the sample size available for each tumour type (23 MB, 29 PA and 19 EP). As discussed in section 3.9, this criterion is useful in reducing the likelihood of over-fitting.

Table 8.6: The optimal feature set for ADC, T2, post-contrast T1 and the combined multimodal MRI based TA. The features are selected by the mRMR and SFS method.

Image Types	Textural Features
ADC	$svar(2,0°)$, $savg(5,135°)$, m, $sosv(4,0°)$, $idmn(2,90°)$, $cshad(1,0°)$
T2	$savg(1,0°)$, $Haar(\mathrm{V2})$, $cshad(1{,}90°)$, $entro(1,90°)$, $svar(1{,}90°)$, $cshad(2{,}90°)$
Post-T1	$cprom(1,0°)$, $LGRE(0°)$, $SRLGE(135°)$, $homo(1,135°)$,
Combined ADC, T2 and post-T1	ADC: $svar(2,0°)$, $savg(5,135°)$, m, $sosv(4,0°)$, $idmn(2,90°)$ T2: $entro(1,90°)$, $Haar(\mathrm{V2})$ Post-T1: $RP(90°)$, var, $imc1(2,135^{\circ})$

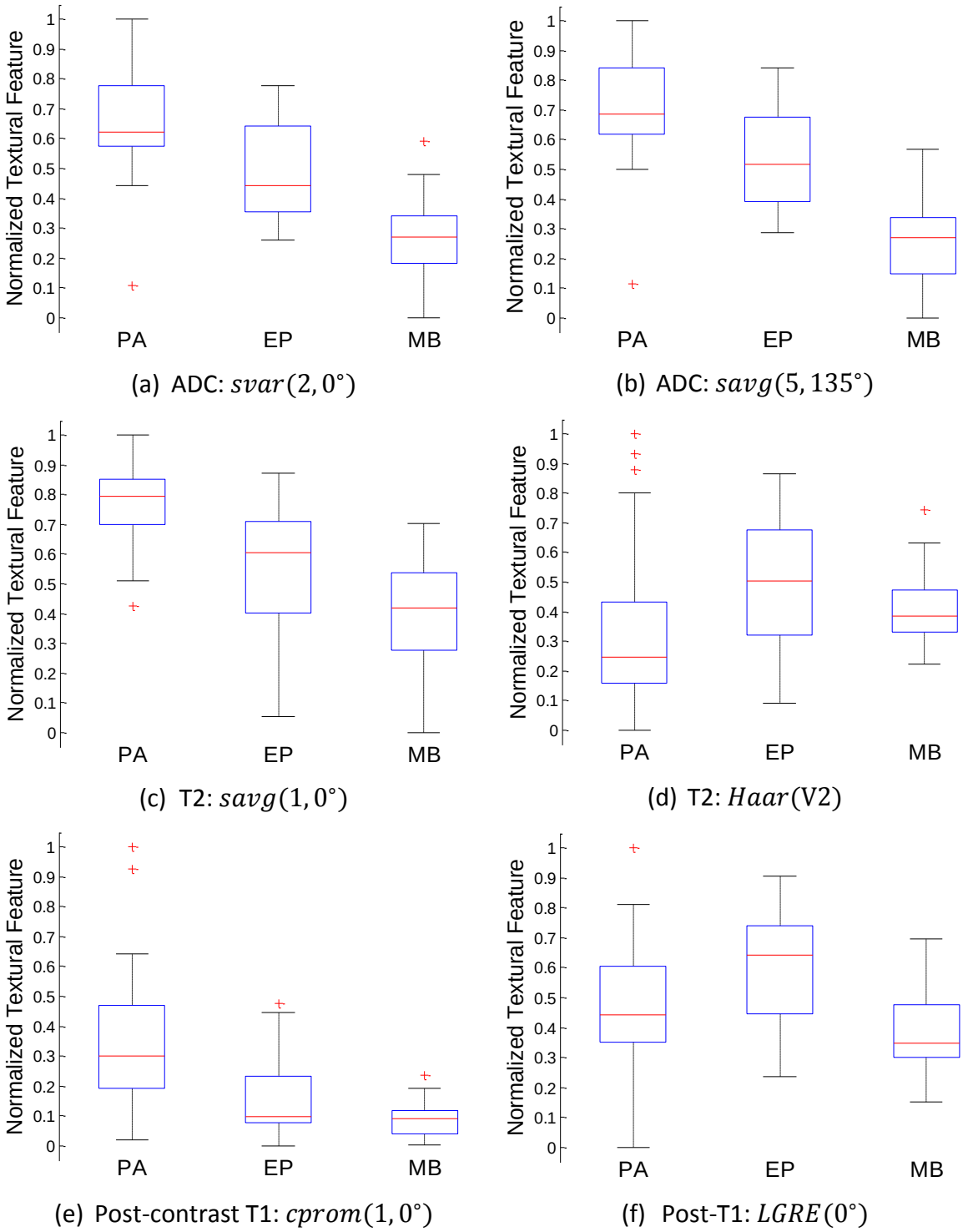


Figure 8.4: Data distribution of the two top-ranked features derived from ADC, T2 and post-contrast T1 dataset. The central red line is the median. The blue edges of the box are the 25th and 75th percentiles. The black whiskers extend to the most extreme data points. The outliers are red marks plotted individually.

In order to examine data distribution of the top-ranked textural features, box plot is employed to visualise the relationship between groups and textural feature values. Data distribution of the two top-ranked features derived from ADC, T2 and post-contrast T1 images (Figure 8.4), do not show absolute separation and has a wide range of overlapping feature values between the different tumour types. This shows that no single textural feature can completely distinguish all three tumour types.

Despite overlapping values, the top-ranked textural features derived from ADC (Figure 8.4 (a) - (b)) shows better separation between tumour types than those derived from T2 and post-contrast T1 (Figure 8.4 (c) - (f)). Furthermore, ADC-based textural features are more clustered within the same group and have fewer outliers.

8.4.2 Classification Outcomes

The confusion matrix[12] is used to analyse the classification results from 1) ADC, 2) T2, 3) post-contrast T1, 4) the combined multimodal MRI based TA, 5) the additive probability outcome, and 6) the superposition of weighted outcomes, as shown in Table 8.7 – 8.12. The classification accuracy derived from each technique is also graphically compared and depicted in Figure 8.5.

From Table 8.7 - 8.12, amongst the single modality MR technique based TA; ADC provides better discrimination of the tumour types than T2 and post-contrast T1. This result corresponds to the finding in chapter 7 of this thesis. In the multimodal MRI based TA, the additive probability outcome approach gives a similar result to ADC based TA, whereas it

[12] The confusion matrix shows the number of predicted classes against actual classes, the classification accuracy and the proportion of predicted members of each tumour type.

is observed to provide lesser accuracy compared to the combined multimodal MRI based TA and the superposition of weighted outcomes. The combined multimodal MRI based TA and the superposition of weighted outcomes gives higher overall accuracy than the ADC, T2 and post-contrast T1 based TA as well as the additive probability outcome approach.

Table 8.7: Confusion matrix for ADC dataset.

Observed	Predicted			
	MB	PA	EP	Accuracy
MB	21	1	1	91.3%
PA	1	27	1	93.1%
EP	3	6	10	52.6%
Overall Percentage	35.2%	47.9%	16.9%	81.7%

Table 8.8: Confusion matrix for T2 dataset.

Observed	Predicted			
	MB	PA	EP	Accuracy
MB	19	0	4	82.6%
PA	1	23	5	79.3%
EP	3	5	11	57.9%
Overall Percentage	32.4%	39.4%	28.2%	74.6%

Table 8.9: Confusion matrix for post-contrast T1 dataset.

Observed	Predicted			
	MB	PA	EP	Accuracy
MB	17	4	2	73.9%
PA	6	21	2	72.4%
EP	5	2	12	63.2%
Overall Percentage	39.4%	38.0%	22.5%	70.4%

Table 8.10: Confusion matrix of the additive probability outcome.

Observed	Predicted			
	MB	PA	EP	Accuracy
MB	21	0	2	91.3%
PA	1	27	1	93.1%
EP	2	7	10	52.6%
Overall Percentage	33.8%	47.9%	18.3%	81.7%

Table 8.11: Confusion matrix of the combined multimodal MRI based TA.

Observed	Predicted			
	MB	PA	EP	Accuracy
MB	23	0	0	100.0%
PA	0	26	3	89.7%
EP	1	3	15	79.0%
Overall Percentage	33.8%	40.8%	25.4%	90.1%

Table 8.12: Confusion matrix of the superposition of weighted outcomes.

Observed	Predicted			
	MB	PA	EP	Accuracy
MB	22	0	1	95.6%
PA	1	27	1	93.1%
EP	2	5	12	63.2%
Overall Percentage	35.2%	45.1%	19.7%	85.9%

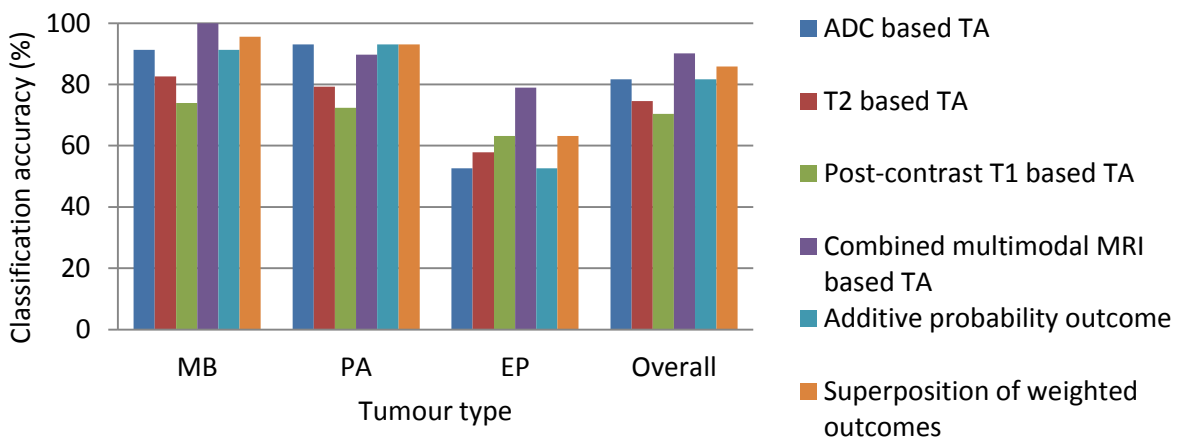


Figure 8.5: Classification accuracy derived from single modality and multimodal MRI based texture analysis. The accuracy figures are measured using LOOCV technique.

8.4.3 Comparative Statistical Test

The statistical hypothesis test of results in section 8.4.2 is shown in Table 8.13. The main conclusions to be drawn from these results are as follows. The classification accuracy obtained from the combined multimodal MRI based TA is significantly better than T2 and post-contrast T1 based TA at 95% confidence interval (p-value = 0.015 and 0.003, respectively). The classification accuracy obtained from the superposition of weighted outcomes is significantly better than post-contrast T1 based TA at 95% confidence interval (p-value = 0.025) and T2 based TA at 90% confidence interval (p-value = 0.092).

Table 8.13: P-values of statistical hypothesis test of pairwise classification comparison. The bold figures are p-values showing significant results at either 95% confidence interval (p-value < 0.05) or 90% confidence interval (p-value < 0.1).

	ADC	T2	Post-T1	Additive probability outcome	Combined multimodal MRI	Superposition of weighted outcome
ADC	1	0.310	0.116	1	0.148	0.494
T2	0.310	1	0.573	0.310	**0.015**	**0.092**
Post-T1	0.116	0.573	1	0.116	**0.003**	**0.025**
Additive probability outcome	1	0.310	0.116	1	0.148	0.494
Combined multimodal MRI	0.148	**0.015**	**0.003**	0.148	1	0.438
Superposition of weighted outcome	0.494	**0.092**	**0.025**	0.494	0.438	1

Next we turn to the comparison of single modality ADC based TA with the combined multimodal MRI based TA and the superposition of weighted outcomes approaches.

Neither approach managed to produce a statistically significant improvement over ADC based TA at a 95% confidence level. However it should be noted that multimodal approaches help to discriminate certain types of tumours where single modality approaches fail. These two techniques are still promising techniques to be examined further with a larger dataset.

The experiment revealed that the combined multimodal MRI based TA is less suitable to be combined in the logistic regression model than the combined single modality MR-based image texture. This is because the difference of textural feature ranges obtained from ADC, T2 and post-contrast T1 can lead to the maximum likelihood not converging within the maximum number of iterations and consequently results in a non-reliable classifier system. In addition, this approach takes excessive computation time because it considers larger combinations of feature sets.

Although the combined multimodal MRI based TA produces higher classification accuracy for overall and ependymoma cases than the superposition of weighted outcomes, the outcomes derived from the two methods are not significantly different at 95% confidence interval. The superposition of weighted outcomes approach also provides improved separation of ependymoma from medulloblastoma and pilocytic astrocytoma and still maintains a good classification for medulloblastoma and pilocytic astrocytoma cases, compared to ADC based TA. As a result, we suggested the combined multimodal MRI based TA and superposition of weighted outcomes as processing frameworks for the classification of childhood brain tumours.

8.4.4 Visual Analysis

In order to visualise the improved results obtained from the combined multimodal MRI based TA and superposition weighted outcomes approach, the cases that were misclassified by ADC based TA but were correctly classified by the suggested techniques are presented. The complicated cases that were misclassified by the two multimodal MRI based TA approaches are also demonstrated to examine the difficulty in classification. Some example images are shown in the following figures and additional images are presented in Appendix D.

Four cases, which were misclassified by ADC based TA, were correctly predicted by the superposition of weighted outcomes approach. Two of these cases are shown in Figure 8.6. It is noticeable that the misclassified cases by ADC based TA have similar appearance as the predicted classes. For example histologically, a case in Figure 8.6 (a) is diagnosed as ependymoma. However, it has a relatively large cystic region compared to the solid portion, which is more typical of a pilocytic astrocytoma and ADC based TA misclassified it as such. Histologically, a case in Figure 8.6 (b) is diagnosed as medulloblastoma but is misclassified as ependymoma. We observe that in this particular case a cystic region is included in the ROI. This has the effect of increasing the mean grey-level intensity within the ROI. Therefore it is likely that the features extracted from ADC images which are sensitive to grey-level intensity caused the classifier to identify this case as an ependymoma, which exhibit high grey-level intensities.

Nine cases, which were misclassified by ADC based TA, were correctly predicted by the combined multimodal MRI based TA approach. Two of these cases are shown in Figure

8.7. Some misclassified cases by ADC based TA have similar appearance as the predicted classes. For example, histologically a case in Figure 8.7 (a) is diagnosed as pilocytic astrocytoma. However, the tumour has a dominant solid portion, which is more typical of a medulloblastoma and ADC based TA misclassified it as such. Histologically a case in Figure 8.7 (b) is diagnosed as anaplastic ependymoma. However it has large cystic component which is also a common feature for a pilocytic astrocytoma and ADC based TA misclassified it as such.

With the superposition of weighted outcomes approach, ten cases were predicted incorrectly. Four of these are displayed in Figure 8.8, showing that these tumours deviate from the typical characteristics for their tumour types. The heterogeneous medulloblastoma was misclassified as ependymoma (Figure 8.8 (a)). The ependymomas with large cystic region were misclassified as pilocytic astrocytoma (Figure 8.8 (b)).

With the combined multimodal MRI based TA approach, seven cases were predicted incorrectly. Four of these cases are displayed in Figure 8.9, showing that these tumours deviate from the typical characteristics for their tumour types. Pilocytic astrocytomas and ependymomas are mostly misclassified. This is because pilocytic astrocytoma with dominant solid portion can be misclassified as ependymoma whereas ependymomas with large cystic region can be misclassified as pilocytic astrocytoma (Figure 8.9 (a)). When anaplastic ependymoma or ependymoma has more homogenous tissue like solid portion, it can be misclassified as medulloblastoma (Figure 8.9 (b)).

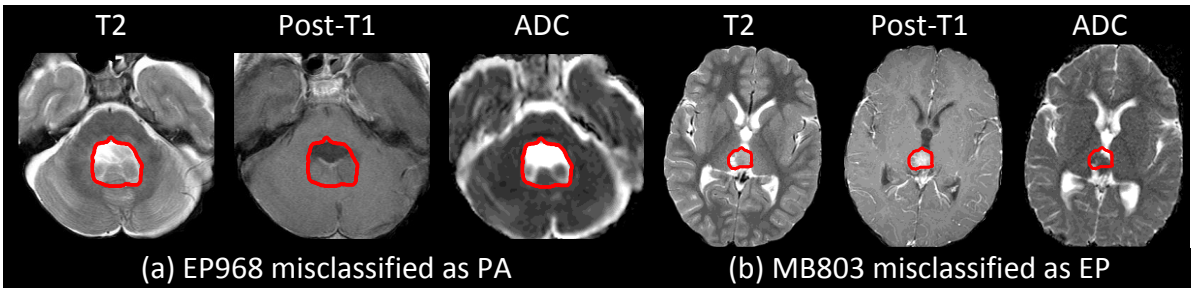


Figure 8.6: Two cases misclassified by ADC based TA but correctly predicted by the superposition of weighted outcomes.

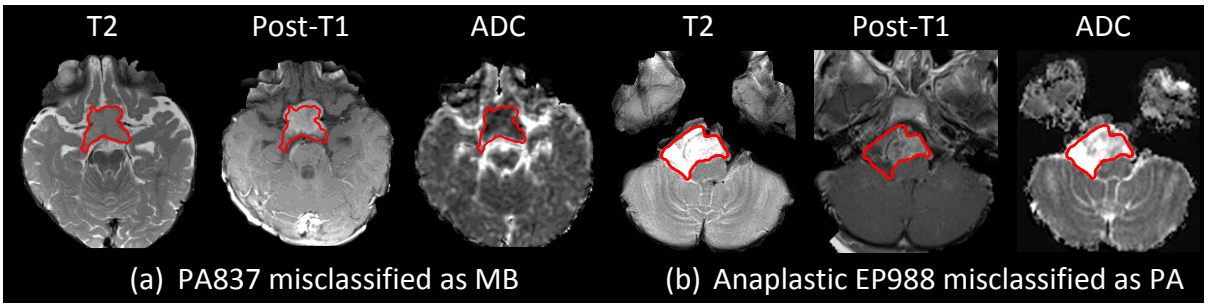


Figure 8.7: Two cases misclassified by ADC based TA but correctly predicted by the combined multimodal MRI based TA approach.

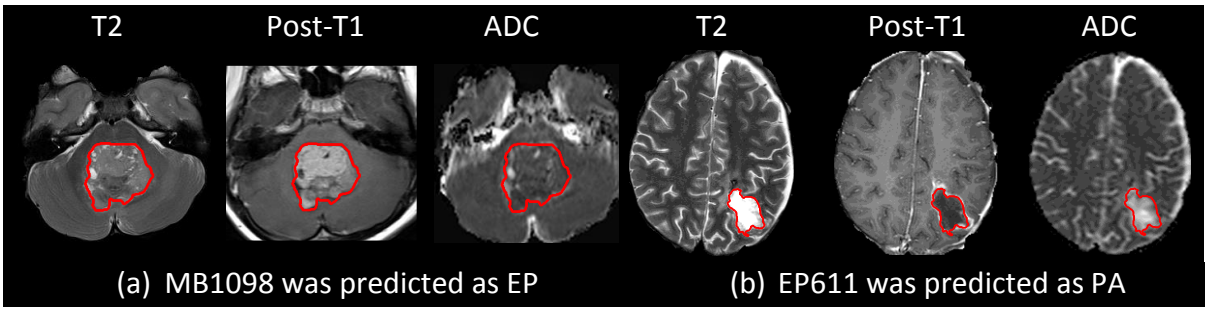


Figure 8.8: Four cases misclassified by the superposition of weighted outcomes.

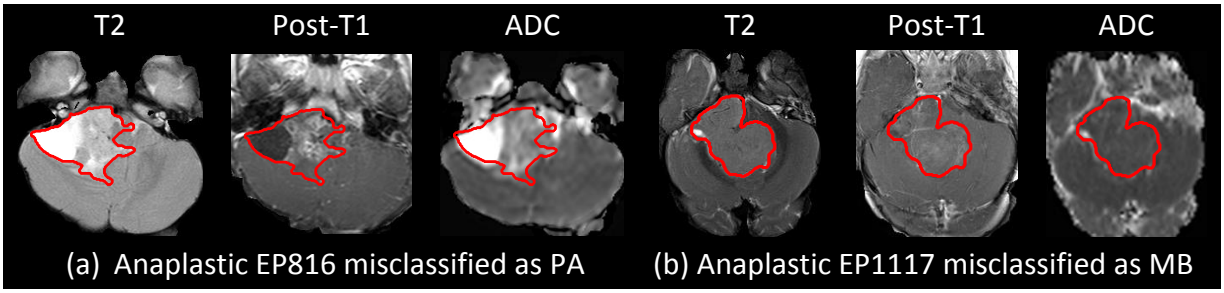


Figure 8.9: Four cases misclassified by the combined multimodal MRI based TA approach.

8.4.5 Building Multinomial Logistic Regression Models

Observing the deviance values, the combined multimodal MRI based TA (deviance = 26) produces a better fitted classification model than the ADC, T2 and post-contrast T1 based TA (deviance = 56, 73 and 87 respectively). However, the combined multimodal MRI based TA does not show significantly better classification accuracy compared to superposition of weighted outcomes approach. As a result, the combined multimodal MRI based TA and the superposition of weighted outcomes approaches are the two promising approaches. We suggested the probability equations of each tumour type can be written by following the concept in equations (3.71) - (3.73), and the multinomial logit models can be written as follows.

Table 8.14: Beta coefficients and goodness of fit derived from ADC dataset. The reference category is ependymoma and the deviance of the model is 55.895.

Predicted class	Variable	B	S.E.	Wald	p-value
Medulloblastoma	Intercept	9.634	6.287	2.348	0.125
	$svar(2, 0°)$	-81.016	30.596	7.012	0.008
	$savg(5, 135°)$	-59.337	57.003	1.084	0.298
	m	-63.574	18.649	11.621	0.001
	$sosv(4, 0°)$	152.192	59.568	6.528	0.012
	$idmn(2,90°)$	11.553	4.592	6.330	0.012
	$cshad(1, 0°)$	10.851	6.759	2.577	0.108
Pilocytic astrocytoma	Intercept	-7.267	3.388	4.601	0.032
	$svar(2, 0°)$	7.042	9.326	0.570	0.450
	$savg(5, 135°)$	34.510	25.904	1.775	0.183
	m	4.218	5.167	0.666	0.414
	$sosv(4, 0°)$	-40.024	25.248	2.513	0.113
	$idmn(2,90°)$	6.093	2.862	4.532	0.033
	$cshad(1, 0°)$	-6.666	5.239	1.619	0.203

The multinomial logit models of ADC dataset are:

$logit\big(P(Y_i = MB)\big) = ln\left(\frac{P(Y_i=MB)}{P(Y_i=EP)}\right)$ = 9.634 + (-81.016) $svar(2, 0°)$ + (-59.337) $savg(5, 135°)$ + (-63.574) m + (152.192) $sosv(4, 0°)$ + (11.553) $idmn(2,90°)$ + (10.851) $cshad(1, 0°)$

$logit\big(P(Y_i = PA)\big) = ln\left(\frac{P(Y_i=PA)}{P(Y_i=EP)}\right)$ = -7.267 + (7.042) $svar(2, 0°)$ + (34.510) $savg(5, 135°)$ + (4.218) m + (-40.024) $sosv(4, 0°)$ + (6.093) $idmn(2,90°)$ + (-6.666) $cshad(1, 0°)$

Table 8.15: Beta coefficients and goodness of fit derived from T2 dataset. The reference category is ependymoma and the deviance of the model is 72.638.

Predicted class	Variable	B	S.E.	Wald	p-value
Medulloblastoma	Intercept	-15.503	7.667	4.089	0.043
	$savg(1,0°)$	53.340	17.211	9.605	0.002
	$Haar(\mathrm{V2})$	-5.297	2.780	3.631	0.057
	$cshad(1,90°)$	276.811	112.366	6.069	0.014
	$entro(1,90°)$	-15.456	4.374	12.486	0.000
	$svar(1,90°)$	-57.961	19.605	8.741	0.003
	$cshad(2,90°)$	-248.621	106.195	5.481	0.019
Pilocytic astrocytoma	Intercept	13.592	5.761	5.566	0.018
	$savg(1,0°)$	27.409	20.666	1.759	0.185
	$Haar(\mathrm{V2})$	-1.494	1.800	0.689	0.407
	$cshad(1,90°)$	-20.969	82.009	0.065	0.798
	$entro(1,90°)$	-14.219	5.443	6.824	0.009
	$svar(1,90°)$	-27.687	20.763	1.778	0.182
	$cshad(2,90°)$	2.818	82.487	0.001	0.973

The multinomial logit models of T2 dataset are:

$logit\big(P(Y_i = MB)\big) = ln\left(\frac{P(Y_i=MB)}{P(Y_i=EP)}\right)$ = -15.503 + (53.340) $savg(1,0°)$ + (-5.297) $Haar(\mathrm{V2})$ + (279.811) $cshad(1,90°)$ + (-15.456) $entro(1,90°)$ + (-57.961) $svar(1,90°)$ + (-248.621) $cshad(2,90°)$

$logit\big(P(Y_i = PA)\big) = ln\left(\frac{P(Y_i=PA)}{P(Y_i=EP)}\right)$ = 13.592 + (27.409) $savg(1,0°)$ + (-1.494) $Haar(\mathrm{V2})$ + (-20.969) $cshad(1,90°)$ + (-14.219) $entro(1,90°)$ +(-27.687) $svar(1,90°)$ + (2.818) $cshad(2,90°)$

Table 8.16: Beta coefficients and goodness of fit derived from post-contrast T1 dataset. The reference category is ependymoma and the deviance of the model is 86.907.

Predicted class	Variable	B	S.E.	Wald	p-value
Medulloblastoma	Intercept	13.008	4.613	7.952	0.005
	$cprom(1, 0°)$	-2.465	4.638	0.282	0.595
	$LGRE\ 0°$	-15.031	5.380	7.806	0.005
	$SRLGE\ 135°$	-3.395	2.334	2.116	0.146
	$homo(1, 135°)$	-8.319	4.204	3.916	0.048
Pilocytic astrocytoma	Intercept	1.272	4.286	0.088	0.767
	$cprom(1, 0°)$	15.488	4.324	12.830	0.000
	$LGRE\ 0°$	-5.647	5.081	1.235	0.266
	$SRLGE\ 135°$	-10.371	3.967	6.835	0.009
	$homo(1, 135°)$	2.535	3.957	0.410	0.522

The multinomial logit models of post-contrast T1 dataset are:

$logit\big(P(Y_i = MB)\big) = ln\left(\frac{P(Y_i=MB)}{P(Y_i=EP)}\right)$ = 13.008 + (-2.465) $cprom(1, 0°)$ + (-15.031) $LGRE\ 0°$ + (-3.395) $SRLGE\ 135°$ + (-8.319) $homo(1, 135°)$

$logit\big(P(Y_i = PA)\big) = ln\left(\frac{P(Y_i=PA)}{P(Y_i=EP)}\right)$ = 1.272 + (15.488) $cprom(1, 0°)$+ (-5.647) $LGRE\ 0°$+ (-10.371) $SRLGE\ 135°$+ (2.535) $homo(1, 135°)$

Table 8.17: Beta coefficients and goodness of fit derived from the combined multimodal MRI based TA approach. The reference category is ependymoma and the deviance of the model is 25.836.

Predicted class	Variable	B	S.E.	Wald	p-value
Medulloblastoma	Intercept	27.766	8.694	10.200	0.001
	ADC: $svar(2, 0°)$	13.570	27.137	0.250	0.617
	ADC: $savg(5, 135°)$	-166.617	57.295	8.457	0.004
	ADC: m	-24.398	10.592	5.306	0.021
	ADC: $sosv(4, 0°)$	143.749	51.349	7.837	0.005
	ADC: $idmn(2, 90°)$	-4.064	7.490	0.294	0.587
	T2: $entro(1, 90°)$	21.851	8.327	6.886	0.009
	T2: $Haar(\mathrm{V2})$	-30.221	8.861	11.632	0.000
	Post-T1: $RP(90°)$	-21.214	7.480	8.043	0.005
	Post-T1: var	11.801	5.791	4.152	0.042
	Post-T1: $imc1(2, 135^{\circ})$	3.911	4.770	0.672	0.412
Pilocytic astrocytoma	Intercept	3.704	5.573	0.442	0.506
	ADC: $svar(2, 0°)$	35.766	21.988	2.646	0.104
	ADC: $savg(5, 135°)$	-85.978	47.709	3.248	0.072
	ADC: m	24.896	9.798	6.456	0.011
	ADC: $sosv(4, 0°)$	29.655	39.058	0.576	0.448
	ADC: $idmn(2, 90°)$	1.988	5.794	0.118	0.732
	T2: $entro(1, 90°)$	2.180	4.300	0.257	0.612
	T2: $Haar(\mathrm{V2})$	-19.462	6.105	10.164	0.001
	Post-T1: $RP(90°)$	-24.946	7.006	12.678	0.000
	Post-T1: var	25.210	6.734	14.016	0.000
	Post-T1: $imc1(2, 135^{\circ})$	11.119	4.248	6.852	0.009

The multinomial logit models of multimodal MRI based TA approach is:

$logit\big(P(Y_i = MB)\big) = ln\left(\frac{P(Y_i=MB)}{P(Y_i=EP)}\right)$ = 27.766 + (13.570) ADC: $svar(2, 0°)$ + (-166.617) ADC: $savg(5, 135°)$ + (-24.398) ADC: m + (143.749) ADC: $sosv(4, 0°)$ + (-4.064) ADC: $idmn(2, 90°)$ + (21.851) T2: $entro(1, 90°)$ + (-30.221) T2: $Haar(\mathrm{V2})$ + (-21.214) Post-T1: $RP(90°)$ + (11.801) Post-T1: var + (3.911) Post-T1: $imc1(2, 135^{\circ})$

$logit\big(P(Y_i = PA)\big) = ln\left(\frac{P(Y_i=PA)}{P(Y_i=EP)}\right)$ = 3.704 + (35.766) ADC: $svar(2, 0°)$ + (-85.978) ADC: $savg(5, 135°)$ + (24.896) ADC: m + (29.655) ADC: $sosv(4, 0°)$ + (1.988)

ADC:$idmn(2, 90°)$ + (2.180) T2: $entro(1, 90°)$ + (-19.462) T2: $Haar(\mathrm{V2})$ + (-24.946) Post-T1: $RP(90°)$ + (25.210) Post-T1: var + (11.119) Post-T1: $imc1(2, 135^{\circ})$

8.4.6 Recommended Processing Framework

Diagnostic classification accuracy based on recent radiological reports (19) is compared with our results obtained from the combined multimodal MRI based TA and superposition of weighted outcomes (Table 8.18). The figures based on radiological reports are the highest estimates as they excude all those where a diagnosis was not given. These unknown diagnostic figures present 52% medulloblastoma, 43% pilocytic astrocytoma, 20% ependymoma and 45% overall (19). Comparing these classification accuracy, our results obtained from combined multimodal MRI based TA and superposition of weighted outcomes approach produce higher classification rates.

Table 8.18: Classification accuracy based on the analysis of radiological reports and our study.

Tumours	Using radiological reports (%)	Combined multimodal MRI based TA (%)	Superposition of weighted outcomes (%)
overall	67	90	86
Medulloblastoma	73	100	96
Pilocytic astrocytoma	75	90	93
Ependymoma	25	79	63

Our two methods provide gains on overall and each tumour based classification compared to single modality MR image based TA and using the radiological reports. As a result, we suggested the two processing frameworks, as represented in Figure 8.10. However, when using different dataset, one processing framework can be more preferable than the other processing framework.

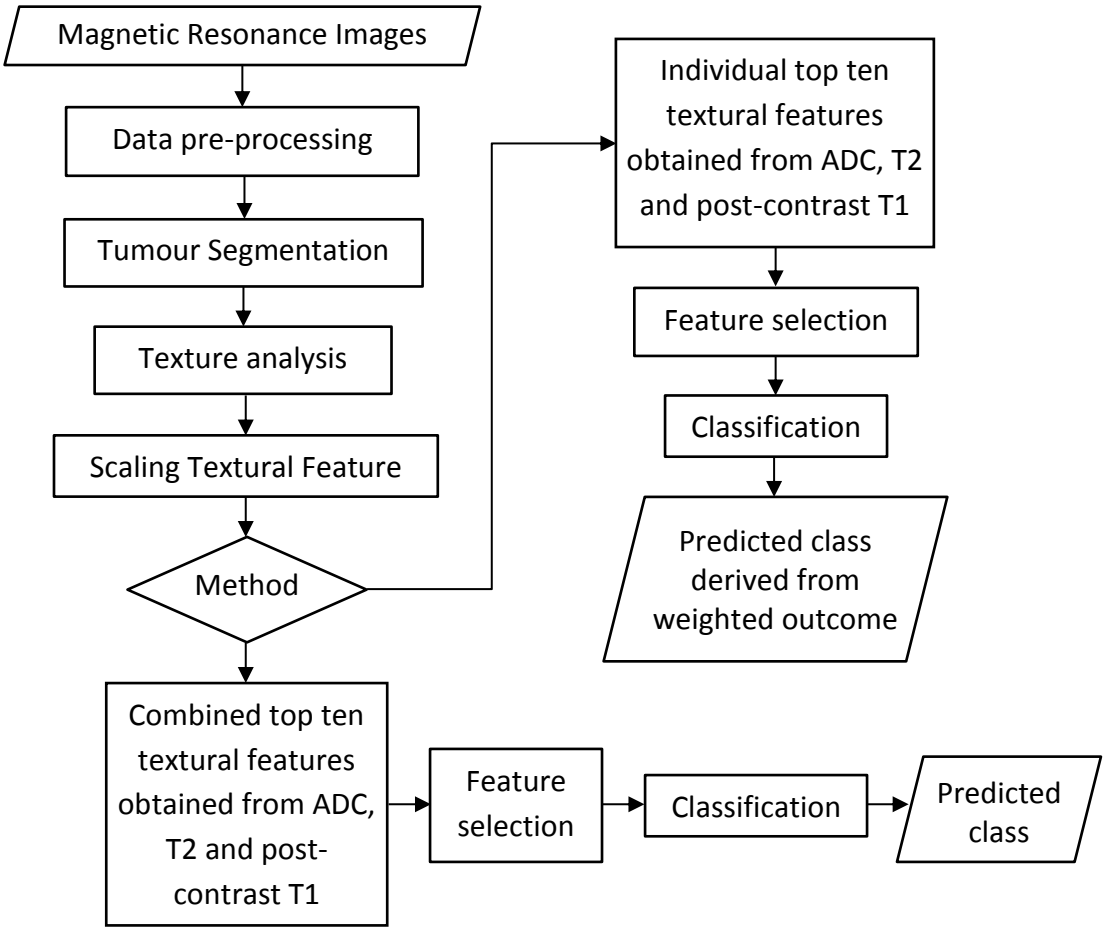


Figure 8.10: The two recommended processing pipelines for the classification of childhood brain tumours. The first pipeline is the combined multimodal MRI based TA. The second pipeline is the superposition weighted outcomes approach.

Andrew Peet, a reader in paediatric oncology at the University of Birmingham and an honorary consultant at Birmingham Children's Hospital suggested that our processing framework is an encouraging study and the results are promising in aiding the clinical management of patients with brain tumours. However, the software analysis package is not approved by regulatory groups, which impedes its clinical use. (Andrew Peet, personal communication, November 27, 2014). To accommodate this issue, there is a need for collaboration between industry, academia and healthcare providers to deliver robust solutions for clinical practices. In addition, appropriate training for radiologists in using software would be beneficial (178).

8.4.7 Biological Perspective

From a biological perspective, tumours can have some similar characteristics resulting in classification difficulties and some distinctive features giving good separable information for the classifiers. The typical imaging features of medulloblastomas are that the tumour is densely packed, small, round to oval, and reduced free water (77, 78). With these imaging characteristics, medulloblastomas, which are highly malignant tumours originating from embryonal cells have promising distinctive textural features from pilocytic astrocytomas, which are slow-growing tumours, and ependymomas, which are mostly composed of solid ependymal cells. The medulloblastomas are reasonably well separated from the other two tumour types.

Ependymomas usually have higher cell density compared to pilocytic astrocytomas, which have low cell density in solid portions with a relatively unrestricted diffusion of water molecules in extracellular space (25). Although these two tumour types are easily

differentiated on the basis of MRI morphology in a majority of cases, diagnosis may be difficult in cases of solid pilocytic astrocytomas without a typical cystic component. In addition, pilocytic astrocytomas and ependymomas located in the cerebellum, near the fourth ventricle, one of the common locations, may have similar MRI appearance (25). As the typical cellularity, ependymomas located in the posterior fossa is somewhere between that of astrocytomas and medulloblastomas (192), the differential diagnosis between ependymomas and the other two tumour types, especially pilocytic astrocytomas, still remains an important question.

8.4.8 Classification Difficulties and Future Work

Various challenges are examined in this thesis, including the selection of region of interest, the selection of study cases with different tumour location and the use of multi-centre dataset. Selection of the ROI is important and mostly obtained by excluding areas of visible cyst, necrosis, haemorrhage, and calcification as much as possible (26, 28, 186, 188, 191, 192, 300). Alteration of ROI coverage varies the value of textural features and the characterisation result. A ROI with large cystic and normal tissues, which may have less informative textural features about histological tumour type, could reduce the overall discriminative power. However, we extracted the largest possible region of abnormal brain tissue by using the semi-automatic segmentation method and derived textural features based on these segments. The textural features derived from such a ROI may be less informative. In addition, atypical characteristics of tumours such as solid pilocytic astrocytoma without typical cystic component and ependymoma with large cystic region were included in the study and resulted in classification difficulties.

The variation of tumour subtypes (e.g. classic medulloblastoma and desmoplastic/nodular medulloblastoma of embryonal tumours as well as ependymoma and anaplastic ependymoma of ependymal tumours), which have biological difference, contributes to the complexity of the classifier. In addition, the variation of cell properties in infratentorial and supratentorial regions also plays an important role in characterising the tumour type. Despite being the same tumour type, cell properties in infratentorial and supratentorial regions can be diverse. Many studies only included tumours in the infratentorial region, while this thesis included tumours from both regions; a majority of cases are located infratentorially and a few cases are located supratentorially. As expected, some of these cases were not correctly classified by the automated methods employed.

A number of previous studies used datasets obtained from a single centre where the same settings are usually applied. However, in this thesis, MR datasets were acquired from a multi-centre setting, where the parameters (e.g. magnetic field, slice thickness, number of scanning direction and b-value) have been configured differently. For example, we use datasets acquired from both 1.5T and 3T scanners. The MRI scans acquired at 3T have higher SNR, greater spatial resolution and reduced motion artefacts because of faster acquisition times compared to the MRI scans at 1.5T (301). As a result, the use of different magnetic field strengths affects the quality of images and hence textural features obtained from those images. In terms of resolution, we use datasets scanned with different slice thickness, which was reported to affect texture details (302). Our datasets have slice thickness across conventional and diffusion MRI scans ranging

between 0.8 to 6 mm. Due to a wide range of slice thickness, textural features derived from these dataset can influence classification performance.

To improve the automatic decision support system, two main approaches within the scope of our study should be investigated: an improved approach of segmentation process and improved methods of TA. As textural features are sensitive to the tissue heterogeneity, the segmentation method should be further developed to separate different tissue types within the region of interest. The current segmentation approach is proposed to extract the largest possible abnormal tumour region. However, the use of solid portion of brain tumours is recommended for classification to improve the differentiation of medulloblastoma, pilocytic astrocytoma and ependymoma (82, 194).

The extraction of tissue subtypes such as cystic, oedema and necrosis tissue from solid tumours is planned for future work. The combination of multimodal MRI was also shown to benefit extraction of distinctive tissue types. This multimodal MR image based segmentation also requires further investigation. In order to improve TA methods, the three approaches were discussed in chapter 7. The number of grey levels, three-dimensional TA and shift-invariant textural features are interesting to be explored.

8.5 Conclusions

This chapter presented the classification of childhood brain tumours based on textural features derived from single modality and multimodal MRI. Two conclusions were drawn.

First, in single modality MRI based TA, ADC based TA provided better differentiation among the three brain tumour types from a multi-scanner multi-centre cohort, compared

with T2- and post-contrast T1-based TA. However, the characterisation of childhood brain tumours can be improved by analysing TA of multimodal MRI. The two approaches: superposition of weighted outcomes and combined multimodal MRI based TA have shown promising results.

Using the superposition of the weighted outcomes approach improves the overall classification accuracy from 82% to 86%, with 10% increased gains for ependymoma, which are particularly difficult to diagnose using TA of single modality MR technique. Using the combined multimodal MRI based TA approach improved the overall classification accuracy from 82% to 90%, with 26% and 9% increased grains for ependymoma and medulloblastoma. However there is minor loss of accuracy (3%) in the case of pilocytic astrocytoma classification.

Second, in clinical application, this approach can be useful to support the diagnostic classification system to identify tumour types. As the system requires a person with no extensive neuroanatomical expertise, it would not be difficult to employ the system in a real clinical practice. However, due to difficulty in obtaining the approval from regulatory bodies, we require the collaboration from industry and healthcare providers to produce a more robust system.

In terms of future work, combining the current finding with other functional MRI such as MRS or perfusion MRI is an interesting area to be explored.

CHAPTER 9 CONCLUSIONS

The aim of this thesis is to develop and examine whether a framework for the texture analysis of multimodal MRI can improve upon the classification accuracy of childhood brain tumours compared with TA of single modality MR technique. To achieve this aim, a multidisciplinary literature review was conducted and a set of focused objectives were defined, researched and evaluated.

The literature review spanned several areas of study, including the principles of MRI and diffusion MRI, MRI characteristics, brain imaging data processing and pattern recognition techniques. All topics were studied with a special emphasis on childhood brain tumours.

The first objective in this thesis is improving the quality of brain tumour segmentation. After evaluating the merits of a number of existing techniques, a novel hybrid semi-automatic segmentation method, based on the MNcut and GVF snake methods is proposed. This combines the advantages of both techniques to enhance the delineation of the tumour boundary. The method was evaluated on four types of structural MRI: T2, FLAIR, pre-contrast T1, and post-contrast T1. T2 based segmentation was scored highest by the experts with consistent inter-rater reliability index and also had high overlap volume ratios and similarity index compared to manual segmentation results. This new segmentation technique would be advantageous in clinical settings, because it is faster than manual segmentation, guarantees reproducibility of results and is less dependent on neuroanatomical expertise.

The second objective is examining tumour classification based on single modality MR textural features in order to understand how differences in image type and feature selection can affect classification accuracy. In order to achieve this objective, experiment 3 was designed where textural features were selected from each image type and presented to the same classifier. Two conclusions were drawn from the results of this experiment. First, ADC, T2 and post-contrast T1 based TA significantly outperformed FLAIR and pre-contrast T1 and tended to produce better discrimination than DTI based TA. The ADC, T2 and post-contrast T1 based TA produce 92%, 92% and 88% overall accuracy respectively. Second, ADC, T2 and post-contrast T1 based TA indicated a good separation between medulloblastoma and pilocytic astrocytoma. However, the differentiation of ependymoma from the other two groups was not satisfactory.

The third objective follows naturally from the observations in the study on single-modality MR technique based TA. That is, whether classification based on multimodal MR textural features can improve upon the single modality results. There are two parts to this objective. First, investigating the methods for multimodal MRI based TA. Second, comparing, the classification results based multimodal and single modality MR textural features. Experiment 4 was designed where the three MR image types: ADC, T2 and post-contrast T1 based TA were integrated into the classification system. The result showed that the classification accuracy could be improved by analysing multimodal MRI using TA. Compared with ADC based TA, the combined multimodal MRI based TA and superposition of the weighted outcomes approach produced 26% and 10% higher diagnostic classification accuracy for ependymoma, which are particularly difficult to diagnose using structural MRI. In addition, compared with using radiological reports, the combined

multimodal MRI based TA and superposition of the weighted outcomes also produce relatively higher classification accuracy.

In terms of future work, certain methods in the processing framework can be improved to enhance performance of the classification system. The segmentation of solid tumours, separating cystic region from region of interest, discussed in chapter 6, 7 and 8, are promising approach to provide better classification outcome. Other procedures such as intensity normalisation discussed in chapter 5 as well as three-dimensional TA and shift-invariant texture matrices discussed in chapter 7, could provide more comprehensive textural features. In addition, combining conventional and diffusion MRI based TA with the metabolite profiles of MRS are an interesting area to be explored. Apart from these, we should seek collaboration between healthcare providers and industries to produce a robust software package and deliver to a real clinical practice.

APPENDIX A: TEXTURAL FEATURES DERIVED FROM GLCM AND GLRLM

A.1 Grey-Level Co-Occurrence Matrix Based Textural Features

For a given image, the GLCM ($\boldsymbol{P}$) and its associated vectors (202, 203, 303) are mathematically defined as follows:

$P(i,j)$ is the $(i,j)th$ element of a normalised grey-tone spatial dependence matrix $\boldsymbol{P}$.

$p_x(i)$ is the ith element of the marginal probability vector $\boldsymbol{p_x}$, obtained by summing the rows of $\boldsymbol{P}$ and given by equation (A.1).

$$p_x(i) = \sum_{j=1}^{N_g} P(i,j) \tag{A.1}$$

$p_y(j)$ is the jth element of the marginal probability vector $\boldsymbol{p_y}$, obtained by summing the columns of $\boldsymbol{P}$ and given by equation (A.2).

$$p_y(j) = \sum_{i=1}^{N_g} P(i,j) \tag{A.2}$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j) \qquad k = 2,3,\dots 2N_g \tag{A.3}$$
$$i + j = k$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j) \qquad k = 0,1,\dots N_g - 1 \tag{A.4}$$
$$|i - j| = k$$

$$\mu_x = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i \cdot P(i,j) \tag{A.5}$$

$$\mu_y = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} j \cdot P(i,j) \tag{A.6}$$

$$\sigma_x = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 \cdot P(i,j) \tag{A.7}$$

$$\sigma_y = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_y)^2 \cdot P(i,j) \tag{A.8}$$

where N_g is the number of distinct grey-levels in the quantized image; μ_x and μ_y are mean of $\boldsymbol{p_x}$ and $\boldsymbol{p_y}$; σ_x and σ_y are the standard deviations of $\boldsymbol{p_x}$ and $\boldsymbol{p_y}$.

Some of the textural features derived from the GLCMs are related to the first-order statistical texture measures. The commonly used 20 textural features from GLCM are adapted in the thesis, as mathematically described below.

Autocorrelation

$$autoc = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij) \cdot P(i,j) \tag{A.9}$$

Autocorrelation describes regularity, fineness and coarseness of the texture. The small primitives

repeated periodically in a short distance have higher $autoc$ value than that in a larger distance. The bigger primitive have decreasing $autoc$ value with increasing distance (212).

Correlation
$$corr = \frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(ij)\cdot P(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y} \tag{A.10}$$

Correlation measures how a pixel is correlated to its neighbour over the ROI. Values of $corr$ ranges between -1 and 1; -1 and 1 refers to a perfectly negatively or positively correlated pixel pairs. NaN (Not a Number) indicates a constant image.

Contrast
$$contr = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N_g}\sum_{j=1}^{N_g} P(i,j) \right\} \tag{A.11}$$

Contrast explains the amount of local variation between a pixel and its neighbour over the ROI. Contrast is 0 for a constant image.

Cluster Prominence
$$cprom = \sum_{i}^{N_g}\sum_{j}^{N_g}\left(i+j-\mu_x-\mu_y\right)^4 P(i,j) \tag{A.12}$$

Cluster prominence measures a peak of data distribution, i.e. measures a kurtosis of the GLCM. A low value of $cprom$ refers to small variation in grey-level (304).

Cluster Shade
$$cshad = \sum_{i}^{N_g}\sum_{j}^{N_g}\left(i+j-\mu_x-\mu_y\right)^3 P(i,j) \tag{A.13}$$

Cluster shade measures the symmetry of data distribution, i.e. measures the skewness of the GLCM. A high value of $cshad$ indicates asymmetry of the GLCM (304).

Dissimilarity
$$dissi = \sum_{i}^{N_g}\sum_{j}^{N_g}|i-j|\cdot P(i,j) \tag{A.14}$$

Dissimilarity measures difference between grey-level intensity. A value of $dissi$ is zero if an image is constant.

Energy (Angular Second Moment)
$$ener = \sum_{i}^{N_g}\sum_{j}^{N_g} P(i,j)^2 \tag{A.15}$$

Energy describes uniformity of texture and has values ranges between 0 and 1. A high $ener$ value represents an image having the constant distribution or a periodic pattern (305).

Entropy
$$entro = -\sum_{i}^{N_g}\sum_{j}^{N_g} P(i,j)\log(P(i,j)) \tag{A.16}$$

Entropy measures the randomness of the grey-level distribution and is inversely correlated to energy. A higher $entro$ value denotes heterogeneous texture.

Homogeneity
$$homo = \sum_{i}^{N_g}\sum_{j}^{N_g}\frac{1}{1+(i-j)^2}P(i,j) \tag{A.17}$$

Homogeneity measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Values of $homo$ range between 0 and 1. Value of $homo$ is 1 for a diagonal GLCM.

Maximum Probability $$maxpr = \underset{i,j}{MAX}\, P(i,j) \tag{A.18}$$

Maximum probability determines the probability of the most predominant pixel pair within an ROI.

Sum of Squares (Variance) $$sosv = \sum_{i}^{N_g}\sum_{j}^{N_g}(i-\mu)^2 \cdot P(i,j) \tag{A.19}$$

Sum of square measures the distribution of GLCM. A higher value of $sosv$ implies the higher heterogeneity of texture. A value of $sosv$ is correlated with the GLCM contrast and the first-order standard deviation value. But when texture has low spatial frequencies and low GLCM contrast, $sosv$ can have a low or high value (305).

Sum Average $$savg = \sum_{i=2}^{2N_g} i\, p_{x+y}(i) \tag{A.20}$$

Sum average measures mean of $\boldsymbol{p_{x+y}}$.

Sum Variance $$svar = \sum_{i=2}^{2N_g}(i - f_{sum\ entropy})^2\, p_{x+y}(i) \tag{A.21}$$

Sum variance measures variance of $\boldsymbol{p_{x+y}}$.

Sum Entropy $$sent = -\sum_{i=2}^{2N_g} p_{x+y}(i)\log\{p_{x+y}(i)\} \tag{A.22}$$

Sum entropy measures entropy of $\boldsymbol{p_{x+y}}$.

Difference Variance $$dvar = variance\ of\ \boldsymbol{p_{x-y}} \tag{A.23}$$

Difference variance measures variance of $\boldsymbol{p_{x-y}}$.

Difference Entropy $$dent = -\sum_{i=0}^{N_{g-1}} p_{x-y}(i)\log\{p_{x-y}(i)\} \tag{A.24}$$

Difference entropy measures entropy of $\boldsymbol{p_{x-y}}$.

Information Measure of Correlation 1 $$imc1 = \frac{HXY - HXY1}{\max\ \{HX, HY\}} \tag{A.25}$$

Information Measure of Correlation 2 $$imc2 = (1 - \exp[-2.0(HXY2 - HXY)])^{1/2} \tag{A.26}$$

$$HXY = -\sum_{i}\sum_{j} P(i,j)\log(P(i,j)) \tag{A.27}$$

$$HXY1 = -\sum_{i}\sum_{j} P(i,j)\log\{p_x(i)p_y(i)\} \tag{A.28}$$

$$HXY2 = -\sum_{i}\sum_{j} p_x(i)p_y(i)\log\{p_x(i)p_y(i)\} \tag{A.29}$$

where HX and HY are entropies of $\boldsymbol{p_x}$ and $\boldsymbol{p_y}$.

Information measure of correlation (306).

Inverse difference normalised

$$idn = \sum_{i}^{N_g}\sum_{j}^{N_g}\frac{1}{1+(i-j)^2/N_g}P(i,j) \quad \text{(A.30)}$$

Inverse difference normalised refers to smoothness of the surface.

Inverse difference moment normalised

$$idmn = \sum_{i}^{N_g}\sum_{j}^{N_g}\frac{1}{1+\left[\frac{i-j}{N_g}\right]^2}P(i,j) \quad \text{(A.31)}$$

Inverse difference moment normalised is inversely related to both contrast and energy. However, this is relevant when one of the parameter is kept constant, for example $idmn$ increases and contrast decrease when energy is constant (305).

A.2 Grey-Level Run-Length Matrix Based Textural Features

For a given image, a GLRLM ($\boldsymbol{P}$) is derived with elements of $P(i,j)$ filling along the row of grey-level i and the column of run-length j. The matrix $\boldsymbol{P}$ can be obtained from four directions of 0°, 45°, 90° and 135°. Let N_g be the number of grey-levels, N_r be the maximum run length, n_r be the total number of runs, n_p be the number of pixels in the image. The associated matrix and vectors derived from the matrix $\boldsymbol{P}$ are defined as follows:

Grey-level run-length pixel number matrix

$$p_p(i,j) = P(i,j)\cdot j \quad \text{(A.32)}$$

Grey-level run-number Vector

$$p_g(i) = \sum_{j=1}^{N_r}P(i,j) \quad \text{(A.33)}$$

Run-length run-number vector

$$p_r(j) = \sum_{i=1}^{N_g}P(i,j) \quad \text{(A.34)}$$

Grey-level run-length-one vector

$$p_o(i) = P(i,1) \quad \text{(A.35)}$$

Textural features obtained from the aforementioned matrix and vectors; $\boldsymbol{P}$, $\boldsymbol{p_p}$, $\boldsymbol{p_g}$, $\boldsymbol{p_r}$ and $\boldsymbol{p_o}$, are mathematically defined as follows:

Short Run Emphasis

$$SRE = \frac{1}{n_r}\sum_{j=1}^{N_r}\frac{p_r(j)}{j^2} \quad \text{(A.36)}$$

Short run emphasis describes the emphasis of short runs of grey-level intensity by dividing with number of run j^2.

Long Run Emphasis

$$LRE = \frac{1}{n_r}\sum_{j=1}^{N_r}p_r(j)\cdot j^2 \quad \text{(A.37)}$$

Long run emphasis describes the emphasis of long runs of grey-level intensity by multiplying with number of run j^2.

Grey-Level Nonuniformity

$$GLN = \frac{1}{n_r}\sum_{i=1}^{N_g}p_g(i)^2 \quad \text{(A.38)}$$

Grey-level non uniformity is proportional with large run length values that are uniformly

distributed.

Run Length Nonuniformity

$$RLN = \frac{1}{n_r}\sum_{j=1}^{N_r} p_r(i)^2 \tag{A.39}$$

Run length nonuniformity encodes long runs that are non-uniformly distributed.

Run Percentage

$$RP = \frac{n_r}{n_p} \tag{A.40}$$

A low value of RP indicates the low variability.

Low Grey-Level Run Emphasis

$$LGRE = \frac{1}{n_r}\sum_{i=1}^{N_g} \frac{p_g(j)}{i^2} \tag{A.41}$$

Low grey-level run emphasis describes the emphasis of runs of low grey-level intensity by dividing grey-level i^2.

High Grey-Level Run Emphasis

$$HGRE = \frac{1}{n_r}\sum_{i=1}^{N_g} p_g(i) \cdot i^2 \tag{A.42}$$

High grey-level run emphasis describes the emphasis of runs of high grey-level intensity by multiplying grey-level i^2.

Short Run Low Grey-Level Emphasis

$$SRLGE = \frac{1}{n_r}\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} \frac{P(i,j)}{i^2 \cdot j^2} \tag{A.43}$$

Short run low grey-level describes the emphasis of short run of low grey-level by dividing grey-level $i^2$and number of run j^2.

Short Run High Grey-Level Emphasis

$$SRHGE = \frac{1}{n_r}\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} \frac{P(i,j) \cdot i^2}{j^2} \tag{A.44}$$

Short run high grey-level describes the emphasis of short run of high grey-level by multiplying with grey-level i^2 and dividing by number of run j^2.

Long Run Low Grey-Level Emphasis

$$LRLGE = \frac{1}{n_r}\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} \frac{P(i,j) \cdot j^2}{i^2} \tag{A.45}$$

Long run low grey-level describes the emphasis of long run of low grey-level by multiplying with number of run j^2 and dividing by grey-level i^2.

Long Run High Grey-Level Emphasis

$$LRHGE = \frac{1}{n_r}\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} P(i,j) \cdot i^2 \cdot j^2 \tag{A.46}$$

Long Run High Grey-Level describes the emphasis of long run of high grey-level by multiplying with number of run j^2 and grey-level i^2.

APPENDIX B: PRE-PROCESSING SHELL SCRIPT

The pre-processing analysis for MRI is operated by using the FSL library. A shell script is written for each study case, as exemplified below.

```
# ------------------------------- Eddy Current Correction for Diffusion MR Images ---------------
eddy_correct  dti_in  dti_ed  0
eddy_correct  dwi_in  dwi_ed  0
# ----------------------------------------------------Skull Stripping ----------------------------------------
bet  T2_in  T2_brain  -f  0.35  -g  0
bet  FLAIR_in  FLAIR_brain  -f  0.3  -g  0
bet  T1pre_in  T1pre_brain  -f  0.35  -g  0
bet  T1post_in  T1post_brain  -f  0.35  -g  0
bet  dti_ed  dti_brain  -f  0.25  -g  0  -m
# ------------------------------ DTI map reconstruction – create FA and MD map -----------------
dtifit  --data=dti_brain  --out=dti  --mask=dti_brain_mask  --bvecs=bvec  --bvals=bval
# ---------------------------------------------- Create ADC maps----------------------------------------
fslroi  dwi_ed  b0  0  128  0  128  0  27  0  1
bet  b0  b0_brain  -f  0.25  -g  0
fslroi  dwi_ed  b1000  0  128  0  128  0  27  1  1
bet  b1000  b1000_brain  -f  0.25  -g  0
fslmaths  b1000_brain  -div  b0_brain  -log  -div  -1000  ADC
# ---------------------------------------------- Registration ---------------------------------------------
flirt  -in  ADC  -ref  T2_brain  -out  T2_ADC  -omat  T2_ADC.mat  -bins  256  -cost  normmi -searchrx  -90  90  -searchry  -90  90  -searchrz  -90  90  -dof  6  -interp  sinc  -sincwidth  7 -sincwindow  hanning

flirt  -in  dti_MD  -ref  T2_brain  -out  T2_MD  -omat  T2_MD.mat  -bins  256  -cost  normmi -searchrx  -90  90  -searchry  -90  90  -searchrz  -90  90  -dof  6  -interp  sinc  -sincwidth  7 -sincwindow  hanning

flirt  -in  dti_FA  -ref  T2_brain  -out  T2_FA  -omat  T2_FA.mat  -bins  256  -cost  normmi -searchrx  -90  90  -searchry  -90  90  -searchrz  -90  90  -dof  6  -interp  sinc  -sincwidth  7 -sincwindow  hanning

flirt  -in  T1pre_brain  -ref  T2_brain  -out  T2_T1pre  -omat  T2_T1pre.mat  -bins  256  -cost  normmi  -searchrx  -90  90  -searchry  -90  90  -searchrz  -90  90  -dof  6  -interp  sinc  -sincwidth  7  -sincwindow  hanning

flirt  -in  T1post_brain  -ref  T2_brain  -out  T2_T1post  -omat  T2_T1post.mat  -bins  256  -cost  normmi  -searchrx  -90  90  -searchry  -90  90  -searchrz  -90  90  -dof  6  -interp  sinc  -sincwidth  7  -sincwindow  hanning

flirt  -in  FLAIR_brain  -ref  T2_brain  -out  T2_FLAIR  -omat  T2_FLAIR.mat  -bins  256  -cost  normmi  -searchrx  -90  90  -searchry  -90  90  -searchrz  -90  90  -dof  6  -interp  sinc  -sincwidth  7  -sincwindow  hanning
```

```
# ---------------------------------------------Linear Intensity Normalisation ---------------------------
scale=4095
bet T2 bet -f 0.2 -m -n
read min max <<< $(fslstats T2_brain -k bet_mask -r)
fslmaths T2_brain -thr $max -bin -mul $scale peaks
fslmaths T2_brain -sub $min -thr 0 -mul $scale -div $(echo $max - $min | /usr/bin/bc ) -uthr
$scale -add peaks T2_norm

bet T2_FLAIR bet -f 0.2 -m -n
read min max <<< $(fslstats T2_FLAIR -k bet_mask -r)
fslmaths T2_FLAIR -thr $max -bin -mul $scale peaks
fslmaths T2_FLAIR -sub $min -thr 0 -mul $scale -div $(echo $max - $min | /usr/bin/bc ) -uthr
$scale -add peaks T2_FLAIR_norm
bet T2_T1pre bet -f 0.2 -m -n
read min max <<< $(fslstats T2_T1pre -k bet_mask -r)
fslmaths T2_T1pre -thr $max -bin -mul $scale peaks
fslmaths T2_T1pre -sub $min -thr 0 -mul $scale -div $(echo $max - $min | /usr/bin/bc ) -uthr
$scale -add peaks T2_T1pre_norm

bet T2_T1post bet -f 0.2 -m -n
read min max <<< $(fslstats T2_T1post -k bet_mask -r)
fslmaths T2_T1post -thr $max -bin -mul $scale peaks
fslmaths T2_T1post -sub $min -thr 0 -mul $scale -div $(echo $max - $min | /usr/bin/bc) -
uthr $scale -add peaks T2_T1post_norm

scale=1
bet T2_ADC bet -f 0.2 -m -n
read min max <<< $(fslstats T2_ADC -k bet_mask -r)
fslmaths T2_ADC -thr $max -bin -mul $scale peaks
fslmaths T2_ADC -sub $min -thr 0 -mul $scale -div $(echo $max - $min | /usr/bin/bc ) -uthr
$scale -add peaks T2_ADC_norm

bet T2_MD bet -f 0.2 -m -n
read min max <<< $(fslstats T2_MD -k bet_mask -r)
fslmaths T2_MD -thr $max -bin -mul $scale peaks
fslmaths T2_MD -sub $min -thr 0 -mul $scale -div $(echo $max - $min | /usr/bin/bc ) -uthr
$scale -add peaks T2_MD_norm

bet T2_FA bet -f 0.2 -m -n
read min max <<< $(fslstats T2_FA -k bet_mask -r)
fslmaths T2_FA -thr $max -bin -mul $scale peaks
fslmaths T2_FA -sub $min -thr 0 -mul $scale -div $(echo $max - $min | /usr/bin/bc ) -uthr
$scale -add peaks T2_FA_norm
```

APPENDIX C: IMAGING PARAMETERS FOR EXPERIMENTAL DATASET

This section lists imaging parameters of images used in the thesis. The detail of TR, TE, TI, resolution, FOV, slice thickness, number of diffusion direction for DTI and scanners used are listed for the four experiments.

C.1 Imaging Parameters of Datasets for Experiment 1

Table C.1.1: Imaging parameters of ADC for experiment 1.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP611	3925	89	1.8x1.8x5	128x128x25	4	1.5	Philips
EP639	2604	68	0.79x0.79x5	288x288x27	4	3	Philips
EP813	8575	87	0.8x0.8x5	256x256x25	5	1.5	GE
EP821	8575	85	0.9x0.9x5	256x256x37	5	1.5	GE
EP904	2700	96	1.8x1.8x7	128x128x19	5	1.5	Siemens
EP968	3700	108	0.9x0.9x7.25	256x256x20	5	1.5	Siemens
EP535	4813	48	1.8x1.8x4.5	128x128x27	4	3	Philips
EP774	8575	85	0.9x0.9x6	256x256x26	5	1.5	GE
EP815	4321	84	1.8x1.8x5	128x128x27	4	1.5	Philips
EP816	2604	68	0.79x0.79x5	288x288x27	4	3	Philips
EP829	3296	70	1.59x1.59x4	144x144x28	4	3	Philips
EP871	4616	100	1.6x1.6x4.4	144x144x30	4	1.5	Philips
EP875	3296	70	1.6x1.6x4	144x144x28	4	3	Philips
EP902	2700	96	1.8x1.8x7	128x128x19	5	1.5	Siemens
EP 936	3700	108	0.9x0.9x7	256x256x20	5	1.5	Siemens
EP 1111	2700	96	1.8x1.8x8	128x128x19	5	1.5	Siemens
EP1115	2700	96	1.8x1.8x6.5	128x128x19	5	1.5	Siemens
EP1117	2700	96	1.8x1.8x7.5	128x128x19	5	1.5	Siemens
MB532	2486	62	1.8x1.8x5	128x128x27	4	3	Philips
MB536	8000	88.6	0.94x0.94x6.5	256x256x20	5	1.5	GE
MB616	8000	94.6	0.98x0.98x4.4	256x256x30	4	1.5	GE
MB622	5826	58	2x2x3.3	112x112x32	3	3	Phillips
MB629	8575	74.6	0.9x0.9x6	256x256x30	5	1.5	GE
MB673	6200	161	1.8x1.8x6.5	128x128x22	5	1.5	Siemens
MB718	4411	84	1.8x1.8x5	128x128x27	4	1.5	Philips
MB719	4620	100	1.6x1.6x4.4	144x144x30	4	1.5	Philips
MB762	5335	83.6	1.8x1.8x5	128x128x32	4	1.5	Philips
MB783	3086	68	0.8x0.8x4.5	288x288x32	4	3	Philips
MB792	5545	89	1.2x1.2x4	192x192x35	4	1.5	Siemens
MB793	4542	100	1.6x1.6x4.4	144x144x30	4	1.5	Philips
MB796	8000	90	2x2x3.3	112x112x36	3	3	Philips
MB801	5000	93	1.2x1.2x4	192x192x30	4	1.5	Siemens
MB803	8575	84.6	0.9x0.9x5	256x256x33	5	1.5	GE
MB804	4438	84	1.8x1.8x5	128x128x27	4	1.5	Philips
MB820	2604	68	0.8x0.8x5	288x288x27	4	3	Philips
MB827	4452	89	1.8x1.8x4.4	128x128x30	4	1.5	Philips
MB834	5400	94	1.2x1.2x4	192x192x32	4	1.5	Siemens
MB844	7300	97	1.2x1.2x4	192x192x42	4	1.5	Siemens
MB860	8575	86	0.9x0.9x5	256x256x28	5	1.5	GE
MB861	8575	85	0.9x0.9x5	256x256x31	5	1.5	GE
MB862	6700	99	1.2x1.2x4	192x192x38	4	1.5	Siemens
MB867	3296	70	1.6x1.6x4	144x144x28	4	3	Philips
MB929	6000	99	1.2x1.2x4	192x192x34	4	1.5	Siemens

PA571	3600	107	0.9x0.9x7.5	256x256x20	5	1.5	Siemens
PA593	5300	161	1.56x1.56x6.5	104x128x19	5	1.5	Siemens
PA607	8700	110	0.94x0.94x6.5	256x256x20	5	1.5	GE
PA613	4300	89	1.8x1.8x4.4	128x128x27	4	1.5	Philips
PA624	2605	68	0.79x0.79x5	288x288x27	4	3	Philips
PA632	2608	68	0.79x0.79x5	288x288x27	4	3	Philips
PA645	5088	99	1.95x1.95x4.4	128x128x27	4	1.5	Philips
PA655	5026	89	1.8x1.8x4.4	128x128x31	4	1.5	Philips
PA666	5826	58	2x2x3.3	112x112x32	3	3	Philips
PA696	3176	57	2x2x3.3	112x112x32	3	3	Philips
PA722	3350	74	2.4x2.4x5	96x96x24	5	1.5	Philips
PA723	5826	58	2x2x3.3	112x112x32	3	3	Philips
PA728	8575	85	0.9x0.9x6	256x256x24	5	1.5	GE
PA731	3094	68	0.8x0.8x4.4	288x288x32	4	3	Philips
PA741	8575	85	0.9x0.9x6	256x256x23	5	1.5	GE
PA753	2990	62	1.6x1.6x5	144x144x32	4	3	Philips
PA766	3288	70	1.75x1.75x4	128x128x30	4	3	Philips
PA784	2604	68	0.79x0.79x5	288x288x27	4	3	Philips
PA791	8000	85	0.9x0.9x5	256x256x29	5	1.5	GE
PA808	2703	68	0.9x0.9x5	256x256x28	4	3	Philips
PA832	5214	100	1.6x1.6x4.4	144x144x34	4	1.5	Philips
PA837	8000	92	0.98x0.98x4.4	256x256x29	4	1.5	GE
PA846	3296	70	1.6x1.6x4	144x144x28	4	3	Philips
PA856	2903	68	0.9x0.9x5	256x256x30	4	3	Philips
PA863	7736	86	2.2x2.2x2.5	104x108x61	2.5	1.5	Siemens
PA873	7129	89	1.2x1.2x4	192x192x45	4	1.5	Siemens
PA883	3700	108	0.9x0.9x7	256x256x20	5	1.5	Siemens
PA890	3296	70	1.6x1.6x4	144x144x28	4	3	Philips
PA903	2700	96	1.8x1.8x6.5	128x128x19	5	1.5	Siemens
PA912	2700	96	1.8x1.8x6.5	128x128x19	5	1.5	Siemens
PA934	6100	89	1.2x1.2x4	192x192x38	4	1.5	Siemens
PA954	3700	108	0.9x0.9x7.5	256x256x20	5	1.5	Siemens
PA961	3700	108	0.9x0.9x7.5	256x256x20	5	1.5	Siemens

Table C.1.2: Imaging parameters of DTI for experiment 1.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner	No Direction
EP639	5826	58	2x2x3.3	112x112x32	3	3	Philips	16
EP821	12000	88	0.94x0.94x2.5	256x256x54	2.5	1.5	GE	26
EP535	4813	48	1.8x1.8x4.5	128x128x27	4	3	Philips	7
EP815	4321	84	1.8x1.8x5	128x128x27	4	1.5	Philips	7
EP816	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
EP829	3296	70	1.6x1.6x4	144x144x28	4	3	Philips	33
EP871	4497	84	1.8x1.8x5	128x128x27	4	1.5	Philips	7
EP875	3296	70	1.6x1.6x4	144x144x28	4	3	Philips	33
MB532	2486	62	1.8x1.8x5	128x128x27	4.2	3	Philips	7
MB536	7100	88	0.94x0.94x4.4	256x256x23	4	1.5	GE	26
MB622	5826	58	2x2x3.3	112x112x32	3	3	Philips	16
MB718	4411	84	1.8x1.8x5	128x128x27	4	1.5	Philips	7
MB719	4628	66	1.75x1.75x5	128x128x27	4	1.5	Philips	16
MB762	5335	84	1.8x1.8x5	128x128x32	4	1.5	Philips	7
MB783	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
MB792	6510	86	2.2x2.2x2.5	108x104x50	2.5	1.5	Siemens	42
MB796	8000	90	2x2x3.3	112x112x36	3	3	Philips	16
MB801	6510	86	2.2x2.2x2.5	104x108x50	2.5	1.5	Siemens	42
MB804	4438	84	1.8x1.8x5	128x128x27	4	1.5	Philips	7
MB820	5827	58	2x2x3.3	112x112x32	3	3	Philips	16
MB834	7000	86	2.2x2.2x2.5	104x108x55	2.5	1.5	Siemens	42
MB844	7609	86	2.2x2.2x2.5	104x108x60	2.5	1.5	Siemens	42
MB860	12000	101	0.94x0.94x2.5	256x256x58	2.5	1.5	GE	26
MB861	12000	95	0.94x0.94x2.5	256x256x64	2.5	1.5	GE	26
MB862	7400	86	2.2x2.2x2.5	104x108x58	2.5	1.5	Siemens	42
MB867	3296	70	1.6x1.6x4	144x144x28	4	3	Philips	33
MB929	6510	86	2.2x2.2x2.5	104x108x50	2.5	1.5	Siemens	42
PA607	8700	110	0.94x0.94x6.5	256x256x20	5	1.5	GE	7
PA632	3504	62	1.8x1.8x5	128x128x27	4.2	3	Philips	7
PA666	5826	58	2.2x2.2x3.3	112x112x32	3	3	Philips	16

PA696	3176	57	2x2x3.3	112x112x32	3	3	Philips	16
PA722	3350	74	2.4x2.4x5	96x96x24	5	1.5	Philips	33
PA723	5826	58	2x2x3.3	112x112x32	3	3	Philips	16
PA731	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
PA753	2990	62	1.6x1.6x5	144x144x32	4	3	Philips	7
PA766	3288	70	1.75x1.75x4	128x128x30	4	3	Philips	16
PA784	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
PA808	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
PA846	3296	70	1.6x1.6x4	144x144x28	4	3	Philips	33
PA856	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
PA863	7736	86	2x2x2.5	104x108x61	2.5	1.5	Siemens	42
PA890	3296	70	1.6x1.6x4	144x144x28	4	3	Philips	33
PA934	6510	86	2.2x2.2x2.5	108x104x50	2.5	1.5	Siemens	42

Table C.1.3: Imaging parameters of T2 for experiment 1.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP611	5050	100	0.45x0.45x5	512x512x25	4	1.5	Philips
EP639	3000	85	0.45x0.45x4.4	512x512x32	4	3	Philips
EP813	5300	84	0.43x0.43x6	512x512x23	5	1.5	GE
EP821	5100	85	0.45x0.45x4	512x512x46	3	1.5	GE
EP904	6180	115	0.66x0.66x6	240x320x22	4	1.5	Siemens
EP968	4000	89	0.43x0.43x7.25	384x512x20	5	1.5	Siemens
EP535	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
EP774	4220	85	0.45x0.45x4	512x512x38	3	1.5	GE
EP815	6055	100	0.4x0.4x4.4	512x512x30	4	1.5	Philips
EP816	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
EP829	3000	80	0.45x0.45x5	512x512x24	4	3	Philips
EP871	6055	100	0.4x0.4x4.4	512x512x30	4	1.5	Philips
EP875	3000	80	0.45x0.45x5	512x512x28	4	3	Philips
EP902	4920	101	0.57x0.57x5.6	306x384x25	4	1.5	Siemens
EP936	4000	89	0.43x0.43x7.5	384x512x20	5	1.5	Siemens
EP1111	4920	95	0.6x0.6x6.4	306x384x25	4	1.5	Siemens
EP1115	6180	14	0.63x0.63x5.8	240x320x22	4	1.5	Siemens
EP1117	6180	14	0.63x0.63x6.8	240x320x22	4	1.5	Siemens
MB532	3000	80	0.45x.45x5	512x512x27	4	3	Philips
MB536	5060	84	0.47x0.47x6.5	512x512x20	5	1.5	GE
MB616	3760	96	0.49x0.49x4.4	512x512x30	4	1.5	GE
MB622	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
MB629	5040	87	0.45x0.45x4	512x512x43	3	1.5	GE
MB673	4560	105	0.90x0.90x6.5	224x256x22	5	1.5	Siemens
MB718	6459	100	0.41x0.41x4.4	512x512x32	4	1.5	Philips
MB719	6070	100	0.45x0.45x4.4	512x512x30	4	1.5	Philips
MB762	6474	100	0.4x0.4x5	512x512x32	4	1.5	Philips
MB783	3000	85	0.45x0.45x4.5	512x512x32	4	3	Philips
MB792	6980	77	0.45x0.45x3.9	402x448x40	3	1.5	Siemens
MB793	6055	100	0.41x0.41x4.4	512x512x30	4	1.5	Philips
MB796	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
MB801	7183	77	0.45x0.45x3.9	402x448x44	3	1.5	Siemens
MB803	4660	85	0.45x0.45x4	512x512x41	3	1.5	GE
MB804	6055	100	0.4x0.4x4.4	512x512x30	4	1.5	Philips
MB820	3000	85	0.45x0.45x4.5	512x512x33	4	3	Philips
MB827	6059	100	0.45x0.45x4.4	512x512x30	4	1.5	Philips
MB834	5230	107	0.72x0.72x5.2	256x320x30	4	1.5	Siemens
MB844	6890	77	0.51x0.51x3.9	402x448x42	3	1.5	Siemens
MB860	4260	86	0.45x0.45x4	512x512x37	3	1.5	GE
MB861	4440	85	0.45x0.45x4	512x512x39	3	1.5	GE
MB862	6980	77	0.45x0.45x3.9	360x448x40	3	1.5	Siemens
MB867	3000	80	0.45x0.45x5	512x512x28	4	3	Philips
MB929	6980	77	0.51x0.51x3.9	402x448x40	3	1.5	Siemens
PA571	4000	89	0.4x0.4x7.5	384x512x20	5	1.5	Siemens
PA593	4970	105	0.4x0.4x6.5	408x512x24	5	1.5	Siemens
PA607	5320	89	0.47x0.47x6.5	512x512x20	5	1.5	GE
PA613	5453	100	0.45x0.45x4.4	512x512x27	4	1.5	Philips
PA624	3000	80	0.44x0.44x5	480x480x27	4	3	Philips
PA632	3000	80	0.45x0.45x5	512x512x27	4	3	Philips
PA645	5453	100	0.45x0.45x4.4	512x512x27	4	1.5	Philips

PA655	6261	100	0.45x0.45x4.4	512x512x31	4	1.5	Philips
PA666	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	3000	85	0.45x0.45x4.5	512x512x22	4	3	Philips
PA722	5595	110	0.45x0.45x6	512x512x25	5	1.5	Philips
PA723	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
PA728	9140	86	0.45x0.45x4	512x512x41	3	1.5	GE
PA731	3000	80	0.45x0.45x4.4	512x512x34	4	3	Philips
PA741	2660	84	0.4x0.4x6	512x512x23	5	1.5	GE
PA753	3000	80	0.45x0.45x5	512x512x27	4	3	Philips
PA766	3000	80	0.45x0.45x5	512x512x28	4	3	Philips
PA784	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA791	4440	85	0.45x0.45x4	512x512x39	3	1.5	GE
PA808	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA832	6863	100	0.4x0.4x4.4	512x512x34	4	1.5	Philips
PA837	7120	95	0.49x0.49x4.4	512x512x29	4	1.5	GE
PA846	3000	80	0.45x0.45x5	512x512x28	4	3	Philips
PA856	3000	85	0.45x0.45x5	512x512x30	4	3	Philips
PA863	6980	77	0.5x0.5x3.9	402x448x42	3	1.5	Siemens
PA873	7020	77	0.45x0.45x3.9	360x448x43	3	1.5	Siemens
PA883	5550	119	0.4x0.4x8.5	384x512x16	5	1.5	Siemens
PA890	3000	80	0.45x0.45x5	512x512x28	4	3	Philips
PA903	4920	101	0.67x0.57x5.6	306x384x25	4	1.5	Siemens
PA912	4920	101	0.57x0.57x5.6	306x384x25	4	1.5	Siemens
PA934	8162	77	0.5x0.5x3.9	402x448x50	3	1.5	Siemens
PA954	4000	89	0.43x0.43x8	384x512x20	5	1.5	Siemens
PA961	4000	89	0.43x0.43x7.5	384x512x20	5	1.5	Siemens

Table C.1.4: Imaging parameters of FLAIR for experiment 1.

Case	TR	TE	TI	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	11000	125	2800	0.45x0.45x4.4	512x512x32	4	3	Philips
EP813	8002	136	2000	0.4x0.4x6	512x512x23	5	1.5	GE
EP821	8002	136	2000	0.45x0.45x6	512x512x30	5	1.5	GE
EP535	11000	125	2800	0.45x0.45x4.5	512x512x27	4	3	Philips
EP774	8002	136	2000	0.45x0.45x6	512x512x24	5	1.5	GE
EP815	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP816	11000	125	2800	0.45x0.45x4.5	512x512x27	4	3	Philips
EP871	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP902	11000	140	2800	0.58x0.58x6.6	400x400x22	6	1.5	Siemens
MB532	11000	125	2800	1x1x5	256x256x27	4	3	Philips
MB616	9627	131	2400	0.98x0.98x4.4	256x256x30	4	1.5	GE
MB629	8002	139	2000	0.45x0.45x6	512x512x30	5	1.5	GE
MB673	9000	122	2500	0.9x0.9x6.5	204x256x24	5	1.5	Siemens
MB718	11000	100	2500	0.9x0.9x4.4	256x256x32	4	1.5	Philips
MB719	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB762	11000	100	2500	0.9x0.9x5	256x256x32	4	1.5	Philips
MB783	11000	125	2800	0.45x0.45x4.5	512x512x32	4	3	Philips
MB792	9000	89	2500	0.6x0.6x6.5	336x384x24	5	1.5	Siemens
MB793	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB796	11000	125	2800	0.45x0.45x4.5	512x512x30	4	3	Philips
MB801	9000	89	2500	0.6x0.6x6.5	336x384x26	5	1.5	Siemens
MB803	8002	136	2000	0.45x0.45x6.5	512x512x24	5	1.5	GE
MB820	11000	125	2800	0.45x0.45x4.5	512x512x33	4	3	Philips
MB827	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB844	9000	89	2500	0.6x0.6x6.5	336x384x26	5	1.5	Siemens
MB860	8002	135	2000	0.45x0.45x6.5	512x512x24	5	1.5	GE
MB861	8002	135	2000	0.45x0.45x6.5	512x512x24	5	1.5	GE
MB862	9000	89	2500	0.6x0.6x6.5	336x384x23	5	1.5	Siemens
MB929	9000	89	2500	0.6x0.6x6.5	336x384x26	5	1.5	Siemens
PA613	11000	100	2500	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA624	11000	125	2800	0.44x0.44x5	480x480x27	4	3	Philips
PA645	11000	100	2500	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA655	11000	100	2500	0.9x0.9x4.4	256x256x31	4	1.5	Philips
PA666	11000	125	2800	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	10000	125	2800	0.45x0.45x4.5	512x512x33	4	3	Philips
PA723	11000	125	2800	0.45x0.45x4.5	512x512x27	4	3	Philips
PA728	8002	136	2000	0.44x0.44x6	512x512x24	5	1.5	GE

PA731	11000	125	2800	0.45x0.45x4.4	512x512x34	4	3	Philips
PA741	8002	136	2000	0.4x0.4x6	512x512x23	5	1.5	GE
PA753	10000	140	2750	1x1x3	224x224x45	3	3	Philips
PA784	11000	125	2800	0.45x0.45x4.5	512x512x30	4	3	Philips
PA791	8002	136	2000	0.44x0.44x6	512x512x24	5	1.5	GE
PA808	11000	125	2800	0.45x0.45x4.5	512x512x32	4	3	Philips
PA832	11000	100	2500	0.9x0.9x4.4	256x256x34	4	1.5	Philips
PA856	11000	125	2800	0.45x0.45x5	512x512x30	4	3	Philips
PA863	9000	89	2500	0.6x0.6x6.5	336x384x25	5	1.5	Siemens
PA873	9000	89	2500	0.6x0.6x6.5	288x384x27	5	1.5	Siemens
PA934	9000	89	2500	0.6x0.6x5	336x384x40	5	1.5	Siemens

Table C.1.5: Imaging parameters of pre-contrast T1 for experiment 1.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	645	14	0.45x0.45x4.4	512x512x32	4	3	Philips
EP813	660	18	0.43x0.43x6	512x512x23	5	1.5	GE
EP821	420	18	0.45x0.45x6.5	512x512x28	5	1.5	GE
EP904	561	13	0.55x0.55x5.4	288x384x25	4.5	1.5	Siemens
EP968	550	14	0.43x0.43x7.25	384x512x20	5	1.5	Siemens
EP535	574	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP774	360	18	0.45x0.45x6.5	512x512x24	5	1.5	GE
EP815	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP816	565	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP871	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP902	561	13	0.57x0.57x5.4	288x384x25	4.5	1.5	Siemens
EP936	550	14	0.43x0.43x7.5	384x512x20	5	1.5	Siemens
EP1111	561	13	0.6x0.6x6.5	288x384x25	4.5	1.5	Siemens
EP1115	561	13	0.5x0.5x5.4	288x384x25	4.5	1.5	Siemens
EP1117	561	13	0.5x0.5x6	288x384x25	4.5	1.5	Siemens
MB532	425	10	0.45x0.45x5	512x512x27	4	3	Philips
MB536	660	22	0.47x0.47x6.5	512x512x20	5	1.5	GE
MB616	700	14	0.49x0.49x4.4	512x512x30	4	1.5	GE
MB622	571	14	0.45x0.45x4.5	512x512x27	4	3	Philips
MB629	420	18	0.45x0.45x6.5	512x512x27	5	1.5	GE
MB673	588	8.7	0.9x0.9x6.5	192x256x23	5	1.5	Siemens
MB718	373	12	0.9x0.9x4.4	256x256x32	4	1.5	Philips
MB719	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB762	721	12	0.9x0.9x5	256x256x32	4	1.5	Philips
MB783	651	14	0.45x0.45x4.5	512x512x32	4	3	Philips
MB792	613	8.4	0.63x0.63x4.8	288x320x30	4	1.5	Siemens
MB793	665	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB796	612	14	0.45x0.45x4.5	512x512x30	4	3	Philips
MB801	613	8.4	0.63x0.63x4.8	288x320x30	4	1.5	Siemens
MB803	380	18	0.45x0.45x6.5	512x512x26	5	1.5	GE
MB827	675	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB844	400	8.4	0.72x0.72x4.8	288x320x35	4	1.5	Siemens
MB860	660	18	0.45x0.45x6.5	512x512x23	5	1.5	GE
MB861	460	18	0.45x0.45x6.5	512x512x25	5	1.5	GE
MB862	675	8.4	0.63x0.63x4.8	288x320x32	4	1.5	Siemens
MB929	654	8.4	0.7x0.7x4.8	288x320x32	4	1.5	Siemens
PA593	500	8.7	0.39x0.39x6.5	416x512x24	5	1.5	Siemens
PA607	700	23	0.47x0.47x6.5	512x512x20	5	1.5	GE
PA613	603	12	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA624	1400	12	0.44x0.44x5	480x480x27	4	3	Philips
PA645	603	12	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA655	698	12	0.9x0.9x4.4	256x256x31	4	1.5	Philips
PA666	603	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	691	14	0.45x0.45x4.5	512x512x33	4	3	Philips
PA722	677	15	0.9x0.9x6	256x256x25	5	1.5	Philips
PA723	566	14	0.45x0.45x4.5	512x512x27	4	3	Philips
PA728	720	18	0.44x0.44x6.5	512x512x25	5	1.5	GE
PA741	360	18	0.4x0.4x6	512x512x23	5	1.5	GE
PA784	604	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA791	360	18	0.44x0.44x6.5	512x512x24	5	1.5	GE
PA808	604	14	0.45x0.45x4.5	512x512x32	4	3	Philips
PA832	766	12	0.9x0.9x4.4	256x256x34	4	1.5	Philips

PA837	660	14	0.49x0.49x4.4	512x512x29	4	1.5	GE
PA856	530	10	0.45x0.45x5	512x512x30	4	3	Philips
PA863	695	13	0.72x0.72x4.8	288x320x34	4	1.5	Siemens
PA873	716	8.4	0.63x0.63x4.8	272x320x35	4	1.5	Siemens
PA883	504	14	0.39x0.39x7	384x512x20	5	1.5	Siemens
PA903	561	13	0.57x0.57x5.85	288x384x25	4.5	1.5	Siemens
PA934	819	8.4	0.72x0.72x4.8	288x320x40	4	1.5	Siemens
PA961	550	14	0.39x0.39x7.5	448x512x20	5	1.5	Siemens

Table C.1.6: Imaging parameters of post-contrast T1 for experiment 1.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP611	400	15	0.9x0.9x5	256x256x25	4	1.5	Philips
EP639	645	14	0.45x0.45x4.4	512x512x32	4	3	Philips
EP813	500	22	0.43x0.43x6	512x512x23	5	1.5	GE
EP821	500	21	0.45x0.45x6.5	512x512x28	5	1.5	GE
EP904	580	11	0.55x0.55x6.75	288x384x20	4.5	1.5	Siemens
EP968	697	14	0.43x0.43x7.5	384x512x20	5	1.5	Siemens
EP535	574	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP774	500	21	0.45x0.45x5	512x512x30	4	1.5	GE
EP815	591	10	0.9x0.9x4.4	256x256x32	4	1.5	Philips
EP816	565	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP829	8.2	3.77	0.84x0.84x1.85	288x288x60	0.8	3	Philips
EP871	300	10	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP902	597	15	0.58x0.58x6.6	400x400x22	6	1.5	Siemens
EP936	697	14	0.43x0.43x7.5	384x512x20	5	1.5	Siemens
EP1111	580	11	0.6x0.6x7.9	288x384x20	4.5	1.5	Siemens
EP1115	580	11	0.55x0.55x6.75	288x384x20	4.5	1.5	Siemens
EP1117	580	11	0.55x0.55x7.65	288x384x16	4.5	1.5	Siemens
MB532	425	10	0.45x.45x5	512x512x27	4	3	Philips
MB536	660	22	0.47x0.47x6.5	512x512x20	5	1.5	GE
MB616	500	22	0.49x0.49x4.4	512x512x30	4	1.5	GE
MB622	571	14	0.45x0.45x4.5	512x512x27	4	3	Philips
MB629	560	21	0.45x0.45x6.5	512x512x28	5	1.5	GE
MB673	828	17	0.45x0.45x5.2	384x512x30	4	1.5	Siemens
MB718	325	10	0.9x0.9x4.4	256x256x32	4	1.5	Philips
MB719	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB762	594	10	0.9x0.9x5	256x256x32	4	1.5	Philips
MB783	651	14	0.45x0.45x4.5	512x512x32	4	3	Philips
MB792	414	17	0.78x0.78x5	256x208x30	4	1.5	Siemens
MB793	565	10	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB796	612	14	0.45x0.45x4.5	512x512x30	4	3	Philips
MB803	500	21	0.45x0.45x5.5	512x512x29	4	1.5	GE
MB827	675	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB834	414	17	0.72x0.72x5	260x320x30	4	1.5	Siemens
MB844	414	17	0.72x0.72x5	260x320x30	4	1.5	Siemens
MB860	400	21	0.45x0.45x6.5	512x512x23	5	1.5	GE
MB861	540	21	0.45x0.45x5	512x512x32	4	1.5	GE
MB862	414	17	0.72x0.72x5	260x320x30	3	1.5	Siemens
MB929	414	17	0.72x0.72x5	260x320x30	4	1.5	Siemens
PA571	697	14	0.4x0.4x7.5	384x512x20	5	1.5	Siemens
PA593	663	17	0.45x0.45x6.5	384x512x24	5	1.5	Siemens
PA607	700	23	0.47x0.47x6.5	512x512x20	5	1.5	GE
PA613	603	12	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA624	567	14	0.44x0.44x5	480x480x27	4	3	Philips
PA645	603	12	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA655	692	12	0.9x0.9x4.4	256x256x31	4	1.5	Philips
PA666	603	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	250	2.3	0.45x0.45x5	512x512x30	4	3	Philips
PA722	656	15	0.9x0.9x6	256x256x20	5	1.5	Philips
PA723	566	14	0.45x0.45x4.5	512x512x27	4	3	Philips
PA728	560	21	0.45x0.45x5	512x512x33	4	1.5	GE
PA731	689	14	0.45x0.45x4.4	512x512x34	4	3	Philips
PA741	520	22	0.4x0.4x5	512x512x30	4	1.5	GE
PA753	565	14	0.45x0.45x5	512x512x27	4	3	Philips
PA784	604	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA791	540	21	0.45x0.45x5	512x512x32	4	1.5	GE

PA808	646	14	0.45x0.45x4.5	512x512x32	4	3	Philips
PA832	628	10	0.9x0.9x4.4	256x256x34	4	1.5	Philips
PA837	500	22	0.49x0.49x4.4	512x512x29	4	1.5	GE
PA856	530	10	0.45x0.45x5	512x512x30	4	3	Philips
PA863	470	17	0.7x0.7x5	260x320x33	4	1.5	Siemens
PA873	525	17	0.72x0.72x5	260x320x37	4	1.5	Siemens
PA883	697	14	0.4x0.4x7	384x512x20	5	1.5	Siemens
PA903	580	11	0.57x0.57x6.75	288x384x20	4.5	1.5	Siemens
PA912	697	14	0.39x0.39x7.5	384x512x20	5	1.5	Siemens
PA934	497	17	0.72x0.72x5.2	320x320x35	4	1.5	Siemens
PA954	697	14	0.43x0.43x7.5	384x512x20	5	1.5	Siemens
PA961	697	14	0.43x0.43x7.5	384x512x20	5	1.5	Siemens

C.2 Imaging Parameters of Datasets for Experiment 2

Table C.2.1: Imaging parameters of T2 for experiment 2.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	3000	85	0.45x0.45x4.4	512x512x32	4	3	Philips
EP821	5100	85	0.45x0.45x4	512x512x46	3	1.5	GE
EP535	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
EP774	4220	85	0.45x0.45x4	512x512x38	3	1.5	GE
EP815	6055	100	0.4x0.4x4.4	512x512x30	4	1.5	Philips
EP816	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
EP871	6055	100	0.4x0.4x4.4	512x512x30	4	1.5	Philips
EP902	4920	101	0.57x0.57x5.6	306x384x25	4	1.5	Siemens
MB532	3000	80	0.45x.45x5	512x512x27	4	3	Philips
MB616	3760	95.7	0.49x0.49x4.4	512x512x30	4	1.5	GE
MB629	5040	87.4	0.45x0.45x4	512x512x43	3	1.5	GE
MB673	4560	105	0.90x0.90x6.5	224x256x22	5	1.5	Siemens
MB718	6459	100	0.41x0.41x4.4	512x512x32	4	1.5	Philips
MB719	6070	100	0.45x0.45x4.4	512x512x30	4	1.5	Philips
MB762	6474	100	0.4x0.4x5	512x512x32	4	1.5	Philips
MB783	3000	85	0.45x0.45x4.5	512x512x32	4	3	Philips
MB793	6055	100	0.41x0.41x4.4	512x512x30	4	1.5	Philips
MB796	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
MB827	6059	100	0.45x0.45x4.4	512x512x30	4	1.5	Philips
MB860	4260	86	0.45x0.45x4	512x512x37	3	1.5	GE
MB861	4440	85	0.45x0.45x4	512x512x39	3	1.5	GE
MB862	6980	77	0.45x0.45x3.9	360x448x40	3	1.5	Siemens
PA613	5453	100	0.45x0.45x4.4	512x512x27	4	1.5	Philips
PA624	3000	80	0.44x0.44x5	480x480x27	4	3	Philips
PA645	5453	100	0.45x0.45x4.4	512x512x27	4	1.5	Philips
PA666	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	3000	85	0.45x0.45x4.5	512x512x22	4	3	Philips
PA728	9140	86	0.45x0.45x4	512x512x41	3	1.5	GE
PA733	3000	85	0.45x0.45x4.5	512x512xx27	4	3	Philips
PA741	2660	84	0.4x0.4x6	512x512x23	5	1.5	GE
PA776	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA784	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA791	4440	85	0.45x0.45x4	512x512x39	3	1.5	GE
PA808	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA856	3000	85	0.45x0.45x5	512x512x30	4	3	Philips
PA873	7020	77	0.45x0.45x3.9	360x448x43	3	1.5	Siemens
PA934	8162	77	402x448x50	0.5x0.5x3.9	3	1.5	Siemens

Table C.2.2: Imaging parameters of FLAIR for experiment 2.

Case	TR	TE	TI	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	11000	125	2800	0.45x0.45x4.4	512x512x32	4	3	Philips
EP821	8002	136	2000	0.45x0.45x6	512x512x30	5	1.5	GE
EP535	11000	125	2800	0.45x0.45x4.5	512x512x27	4	3	Philips
EP774	8002	136	2000	0.45x0.45x6	512x512x24	5	1.5	GE
EP815	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips

EP816	11000	125	2800	0.45x0.45x4.5	512x512x27	4	3	Philips
EP871	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP902	11000	140	2800	0.58x0.58x6.6	400x400x22	6	1.5	Siemens
MB532	11000	125	2800	1x1x5	256x256x27	4	3	Philips
MB616	9627	131	2400	0.98x0.98x4.4	256x256x30	4	1.5	GE
MB629	8002	139	2000	0.45x0.45x6	512x512x30	5	1.5	GE
MB673	9000	122	2500	0.9x0.9x6.5	204x256x24	5	1.5	Siemens
MB718	11000	100	2500	0.9x0.9x4.4	256x256x32	4	1.5	Philips
MB719	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB762	11000	100	2500	0.9x0.9x5	256x256x32	4	1.5	Philips
MB783	11000	125	2800	0.45x0.45x4.5	512x512x32	4	3	Philips
MB793	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB796	11000	125	2800	0.45x0.45x4.5	512x512x30	4	3	Philips
MB827	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB860	8002	135	2000	0.45x0.45x6.5	512x512x24	5	1.5	GE
MB861	8002	135	2000	0.45x0.45x6.5	512x512x24	5	1.5	GE
MB862	9000	89	2500	0.6x0.6x6.5	336x384x23	5	1.5	Siemens
PA613	11000	100	2500	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA624	11000	125	2800	0.44x0.44x5	480x480x27	4	3	Philips
PA645	11000	100	2500	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA666	11000	125	2800	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	10000	125	2800	0.45x0.45x4.5	512x512x33	4	3	Philips
PA728	8002	136	2000	0.44x0.44x6	512x512x24	5	1.5	GE
PA733	11000	125	2800	0.44x0.44x4.5	512x512x27	4	3	Philips
PA741	8002	136	2000	0.4x0.4x6	512x512x23	5	1.5	GE
PA776	11000	125	2800	0.44x0.44x4.5	512x512x30	4	3	Philips
PA784	11000	125	2800	0.45x0.45x4.5	512x512x30	4	3	Philips
PA791	8002	136	2000	0.44x0.44x6	512x512x24	5	1.5	GE
PA808	11000	125	2800	0.45x0.45x4.5	512x512x32	4	3	Philips
PA856	11000	125	2800	0.45x0.45x5	512x512x30	4	3	Philips
PA873	9000	89	2500	0.6x0.6x6.5	288x384x27	5	1.5	Siemens
PA934	9000	89	2500	0.6x0.6x5	336x384x40	5	1.5	Siemens

Table C.2.3: Imaging parameters of pre-contrast T1 for experiment 2.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	645	14	0.45x0.45x4.4	512x512x32	4	3	Philips
EP821	420	18	0.45x0.45x6.5	512x512x28	5	1.5	GE
EP535	574	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP774	360	18	0.45x0.45x6.5	512x512x24	5	1.5	GE
EP815	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP816	565	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP871	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP902	561	13	0.57x0.57x5.4	288x384x25	4.5	1.5	Siemens
MB532	425	10	0.45x0.45x5	512x512x27	4	3	Philips
MB616	700	14	0.49x0.49x4.4	512x512x30	4	1.5	GE
MB629	420	18	0.45x0.45x6.5	512x512x27	5	1.5	GE
MB673	588	8.7	0.9x0.9x6.5	192x256x23	5	1.5	Siemens
MB718	373	12	0.9x0.9x4.4	256x256x32	4	1.5	Philips
MB719	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB762	721	12	0.9x0.9x5	256x256x32	4	1.5	Philips
MB783	651	14	0.45x0.45x4.5	512x512x32	4	3	Philips
MB793	665	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB796	612	14	0.45x0.45x4.5	512x512x30	4	3	Philips
MB827	675	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB860	660	18	0.45x0.45x6.5	512x512x23	5	1.5	GE
MB861	460	18	0.45x0.45x6.5	512x512x25	5	1.5	GE
MB862	675	8.4	0.63x0.63x4.8	288x320x32	4	1.5	Siemens
PA613	603	12	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA624	1400	12	0.44x0.44x5	480x480x27	4	3	Philips
PA645	603	12	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA666	603	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	691	14	0.45x0.45x4.5	512x512x33	4	3	Philips
PA728	720	18	0.44x0.44x6.5	512x512x25	5	1.5	GE
PA733	566	14	0.44x0.44x4.5	512x512x27	4	3	Philips
PA741	360	18	0.4x0.4x6	512x512x23	5	1.5	GE
PA776	610	14	0.44x0.44x4.5	512x512x30	4	3	Philips

PA784	604	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA791	360	18	0.44x0.44x6.5	512x512x24	5	1.5	GE
PA808	604	14	0.45x0.45x4.5	512x512x32	4	3	Philips
PA856	530	10	0.45x0.45x5	512x512x30	4	3	Philips
PA873	716	8.4	0.63x0.63x4.8	272x320x35	4	1.5	Siemens
PA934	819	8.4	0.72x0.72x4.8	288x320x40	4	1.5	Siemens

Table C.2.4: Imaging parameters of post-contrast T1 for experiment 2.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	645	14	0.45x0.45x4.4	512x512x32	4	3	Philips
EP821	500	21	0.45x0.45x6.5	512x512x28	5	1.5	GE
EP535	574	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP774	500	21	0.45x0.45x5	512x512x30	4	1.5	GE
EP815	591	10	0.9x0.9x4.4	256x256x32	4	1.5	Philips
EP816	565	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP871	300	10	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP902	597	15	0.58x0.58x6.6	400x400x22	6	1.5	Siemens
MB532	425	10	0.45x.45x5	512x512x27	4	3	Philips
MB616	500	22	0.49x0.49x4.4	512x512x30	4	1.5	GE
MB629	560	21	0.45x0.45x6.5	512x512x28	5	1.5	GE
MB673	828	17	0.45x0.45x5.2	384x512x30	4	1.5	Siemens
MB718	325	10	0.9x0.9x4.4	256x256x32	4	1.5	Philips
MB719	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB762	594	10	0.9x0.9x5	256x256x32	4	1.5	Philips
MB783	651	14	0.45x0.45x4.5	512x512x32	4	3	Philips
MB793	565	10	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB796	612	14	0.45x0.45x4.5	512x512x30	4	3	Philips
MB827	675	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB860	400	21	0.45x0.45x6.5	512x512x23	5	1.5	GE
MB861	540	21	0.45x0.45x5	512x512x32	4	1.5	GE
MB862	414	17	0.72x0.72x5	260x320x30	3	1.5	Siemens
PA613	603	12	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA624	567	14	0.44x0.44x5	480x480x27	4	3	Philips
PA645	603	12	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA666	603	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	250	2.3	0.45x0.45x5	512x512x30	4	3	Philips
PA728	560	21	0.45x0.45x5	512x512x33	4	1.5	GE
PA733	566	14	0.45x0.45x4.5	512x512x27	4	3	Philips
PA741	520	22	0.4x0.4x5	512x512x30	4	1.5	GE
PA776	610	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA784	604	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA791	540	21	0.45x0.45x5	512x512x32	4	1.5	GE
PA808	646	14	0.45x0.45x4.5	512x512x32	4	3	Philips
PA856	530	10	0.45x0.45x5	512x512x30	4	3	Philips
PA856	530	10	0.45x0.45x5	512x512x30	4	3	Philips
PA934	497	17	0.72x072x5.2	320x320x35	4	1.5	Siemens

C.3 Imaging Parameters of Datasets for Experiment 3

Table C.3.1: Imaging parameters of ADC for experiment 3.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	2604	68	0.79x0.79x5	288x288x27	4	3	Philips
EP821	8575	85	0.9x0.9x5	256x256x37	5	1.5	GE
EP535	4813	48	1.8x1.8x4.5	128x128x27	4	3	Philips
EP815	4321	84	1.8x1.8x5	128x128x27	4	1.5	Philips
EP816	2604	68	0.79x0.79x5	288x288x27	4	3	Philips
EP871	4616	100	1.6x1.6x4.4	144x144x30	4	1.5	Philips
MB532	2486	62	1.8x1.8x5	128x128x27	4	3	Philips
MB718	4411	84	1.8x1.8x5	128x128x27	4	1.5	Philips
MB719	4620	100	1.6x1.6x4.4	144x144x30	4	1.5	Philips
MB762	5335	83.6	1.8x1.8x5	128x128x32	4	1.5	Philips
MB783	3086	68	0.8x0.8x4.5	288x288x32	4	3	Philips
MB792	5545	89	1.2x1.2x4	192x192x35	4	1.5	Siemens

MB796	8000	90	2x2x3.3	112x112x36	3	3	Philips
MB844	7300	97	1.2x1.2x4	192x192x42	4	1.5	Siemens
MB860	8575	86	0.9x0.9x5	256x256x28	5	1.5	GE
MB861	8575	85	0.9x0.9x5	256x256x31	5	1.5	GE
MB862	6700	99	1.2x1.2x4	192x192x38	4	1.5	Siemens
MB929	6000	99	1.2x1.2x4	192x192x34	4	1.5	Siemens
PA666	5826	58	2x2x3.3	112x112x32	3	3	Philips
PA696	3176	57	2x2x3.3	112x112x32	3	3	Philips
PA723	5826	58	2x2x3.3	112x112x32	3	3	Philips
PA784	2604	68	0.79x0.79x5	288x288x27	4	3	Philips
PA808	2703	68	0.9x0.9x5	256x256x28	4	3	Philips
PA856	2903	68	0.9x0.9x5	256x256x30	4	3	Philips

Table C.3.2: Imaging parameters of DTI for experiment 3.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner	No Direction
EP639	5826	58	2x2x3.3	112x112x32	3	3	Philips	16
EP821	12000	88	0.94x0.94x2.5	256x256x54	2.5	1.5	GE	26
EP535	4813	48	1.8x1.8x4.5	128x128x27	4	3	Philips	7
EP815	4321	84	1.8x1.8x5	128x128x27	4	1.5	Philips	7
EP816	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
EP871	4497	84	1.8x1.8x5	128x128x27	4	1.5	Philips	7
MB532	2486	62	1.8x1.8x5	128x128x27	4.2	3	Philips	7
MB718	4411	84	1.8x1.8x5	128x128x27	4	1.5	Philips	7
MB719	4628	66	1.75x1.75x5	128x128x27	4	1.5	Philips	16
MB762	5335	84	1.8x1.8x5	128x128x32	4	1.5	Philips	7
MB783	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
MB792	6510	86	2.2x2.2x2.5	108x104x50	2.5	1.5	Siemens	42
MB796	8000	90	2x2x3.3	112x112x36	3	3	Philips	16
MB844	7609	86	2.2x2.2x2.5	104x108x60	2.5	1.5	Siemens	42
MB860	12000	101	0.94x0.94x2.5	256x256x58	2.5	1.5	GE	26
MB861	12000	95	0.94x0.94x2.5	256x256x64	2.5	1.5	GE	26
MB862	7400	86	2.2x2.2x2.5	104x108x58	2.5	1.5	Siemens	42
MB929	6510	86	2.2x2.2x2.5	104x108x50	2.5	1.5	Siemens	42
PA666	5826	58	2.2x2.2x3.3	112x112x32	3	3	Philips	16
PA696	3176	57	2x2x3.3	112x112x32	3	3	Philips	16
PA723	5826	58	2x2x3.3	112x112x32	3	3	Philips	16
PA784	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
PA808	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
PA856	8000	90	2x2x3.3	112x112x32	3	3	Philips	16
PA863	7736	86	2x2x2.5	104x108x61	2.5	1.5	Siemens	42
PA934	6510	86	2.2x2.2x2.5	108x104x50	2.5	1.5	Siemens	42

Table C.3.3: Imaging parameters of T2 for experiment 3.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	3000	85	0.45x0.45x4.4	512x512x32	4	3	Philips
EP821	5100	85	0.45x0.45x4	512x512x46	3	1.5	GE
EP535	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
EP815	6055	100	0.4x0.4x4.4	512x512x30	4	1.5	Philips
EP816	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
EP871	6055	100	0.4x0.4x4.4	512x512x30	4	1.5	Philips
MB532	3000	80	0.45x.45x5	512x512x27	4	3	Philips
MB718	6459	100	0.41x0.41x4.4	512x512x32	4	1.5	Philips
MB719	6070	100	0.45x0.45x4.4	512x512x30	4	1.5	Philips
MB762	6474	100	0.4x0.4x5	512x512x32	4	1.5	Philips
MB783	3000	85	0.45x0.45x4.5	512x512x32	4	3	Philips
MB792	6980	77	0.45x0.45x3.9	402x448x40	3	1.5	Siemens
MB796	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
MB844	6890	77	0.51x0.51x3.9	402x448x42	3	1.5	Siemens
MB860	4260	86	0.45x0.45x4	512x512x37	3	1.5	GE
MB861	4440	85	0.45x0.45x4	512x512x39	3	1.5	GE
MB862	6980	77	0.45x0.45x3.9	360x448x40	3	1.5	Siemens
MB929	6980	77	0.51x0.51x3.9	402x448x40	3	1.5	Siemens
PA666	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	3000	85	0.45x0.45x4.5	512x512x22	4	3	Philips

PA723	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
PA784	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA808	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA856	3000	85	0.45x0.45x5	512x512x30	4	3	Philips
PA863	6980	77	0.5x0.5x3.9	402x448x42	3	1.5	Siemens
PA934	8162	77	0.5x0.5x3.9	402x448x50	3	1.5	Siemens

Table C.3.4: Imaging parameters of FLAIR for experiment 3.

Case	TR	TE	TI	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	11000	125	2800	0.45x0.45x4.4	512x512x32	4	3	Philips
EP821	8002	136	2000	0.45x0.45x6	512x512x30	5	1.5	GE
EP535	11000	125	2800	0.45x0.45x4.5	512x512x27	4	3	Philips
EP815	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP816	11000	125	2800	0.45x0.45x4.5	512x512x27	4	3	Philips
EP871	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB532	11000	125	2800	1x1x5	256x256x27	4	3	Philips
MB718	11000	100	2500	0.9x0.9x4.4	256x256x32	4	1.5	Philips
MB719	11000	100	2500	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB762	11000	100	2500	0.9x0.9x5	256x256x32	4	1.5	Philips
MB783	11000	125	2800	0.45x0.45x4.5	512x512x32	4	3	Philips
MB792	9000	89	2500	0.6x0.6x6.5	336x384x24	5	1.5	Siemens
MB796	11000	125	2800	0.45x0.45x4.5	512x512x30	4	3	Philips
MB844	9000	89	2500	0.6x0.6x6.5	336x384x26	5	1.5	Siemens
MB860	8002	135	2000	0.45x0.45x6.5	512x512x24	5	1.5	GE
MB861	8002	135	2000	0.45x0.45x6.5	512x512x24	5	1.5	GE
MB862	9000	89	2500	0.6x0.6x6.5	336x384x23	5	1.5	Siemens
MB929	9000	89	2500	0.6x0.6x6.5	336x384x26	5	1.5	Siemens
PA666	11000	125	2800	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	10000	125	2800	0.45x0.45x4.5	512x512x33	4	3	Philips
PA723	11000	125	2800	0.45x0.45x4.5	512x512x27	4	3	Philips
PA784	11000	125	2800	0.45x0.45x4.5	512x512x30	4	3	Philips
PA808	11000	125	2800	0.45x0.45x4.5	512x512x32	4	3	Philips
PA856	11000	125	2800	0.45x0.45x5	512x512x30	4	3	Philips
PA863	9000	89	2500	0.6x0.6x6.5	336x384x25	5	1.5	Siemens
PA934	9000	89	2500	0.6x0.6x5	336x384x40	5	1.5	Siemens

Table C.3.5: Imaging parameters of pre-contrast T1 for experiment 3.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	645	14	0.45x0.45x4.4	512x512x32	4	3	Philips
EP821	420	18	0.45x0.45x6.5	512x512x28	5	1.5	GE
EP535	574	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP815	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP816	565	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP871	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB532	425	10	0.45x0.45x5	512x512x27	4	3	Philips
MB718	373	12	0.9x0.9x4.4	256x256x32	4	1.5	Philips
MB719	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB762	721	12	0.9x0.9x5	256x256x32	4	1.5	Philips
MB783	651	14	0.45x0.45x4.5	512x512x32	4	3	Philips
MB792	613	8.4	0.63x0.63x4.8	288x320x30	4	1.5	Siemens
MB796	612	14	0.45x0.45x4.5	512x512x30	4	3	Philips
MB844	400	8.4	0.72x0.72x4.8	288x320x35	4	1.5	Siemens
MB860	660	18	0.45x0.45x6.5	512x512x23	5	1.5	GE
MB861	460	18	0.45x0.45x6.5	512x512x25	5	1.5	GE
MB862	675	8.4	0.63x0.63x4.8	288x320x32	4	1.5	Siemens
MB929	654	8.4	0.72x0.72x4.8	288x320x32	4	1.5	Siemens
PA666	603	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	691	14	0.45x0.45x4.5	512x512x33	4	3	Philips
PA723	566	14	0.45x0.45x4.5	512x512x27	4	3	Philips
PA784	604	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA808	604	14	0.45x0.45x4.5	512x512x32	4	3	Philips
PA856	530	10	0.45x0.45x5	512x512x30	4	3	Philips
PA863	695	13	0.72x0.72x4.8	288x320x34	4	1.5	Siemens
PA934	819	8.4	0.72x0.72x4.8	288x320x40	4	1.5	Siemens

Table C.3.6: Imaging parameters of post-contrast T1 for experiment 3.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP639	645	14	0.45x0.45x4.4	512x512x32	4	3	Philips
EP821	500	21	0.45x0.45x6.5	512x512x28	5	1.5	GE
EP535	574	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP815	591	10	0.9x0.9x4.4	256x256x32	4	1.5	Philips
EP816	565	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP871	300	10	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB532	425	10	0.45x.45x5	512x512x27	4	3	Philips
MB718	325	10	0.9x0.9x4.4	256x256x32	4	1.5	Philips
MB719	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB762	594	10	0.9x0.9x5	256x256x32	4	1.5	Philips
MB783	651	14	0.45x0.45x4.5	512x512x32	4	3	Philips
MB792	414	17	0.78x0.78x5	256x208x30	4	1.5	Siemens
MB796	612	14	0.45x0.45x4.5	512x512x30	4	3	Philips
MB844	414	17	0.72x0.72x5	260x320x30	4	1.5	Siemens
MB860	400	21	0.45x0.45x6.5	512x512x23	5	1.5	GE
MB861	540	21	0.45x0.45x5	512x512x32	4	1.5	GE
MB862	414	17	0.72x0.72x5	260x320x30	3	1.5	Siemens
MB929	414	17	0.72x0.72x5	260x320x30	4	1.5	Siemens
PA666	603	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	250	2.3	0.45x0.45x5	512x512x30	4	3	Philips
PA723	566	14	0.45x0.45x4.5	512x512x27	4	3	Philips
PA784	604	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA808	646	14	0.45x0.45x4.5	512x512x32	4	3	Philips
PA856	530	10	0.45x0.45x5	512x512x30	4	3	Philips
PA863	470	17	0.7x0.7x5	260x320x33	4	1.5	Siemens
PA934	497	17	0.72x072x5.2	320x320x35	4	1.5	Siemens

C.4 Imaging Parameters of Datasets for Experiment 4

Table C.4.1: Imaging parameters of ADC maps for experiment 4.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP611	3925	89	1.8x1.8x5	128x128x25	4	1.5	Philips
EP639	2604	68	0.79x0.79x5	288x288x27	4	3	Philips
EP813	8575	87	0.8x0.8x5	256x256x25	5	1.5	GE
EP821	8575	85	0.9x0.9x5	256x256x37	5	1.5	GE
EP904	2700	96	1.8x1.8x7	128x128x19	5	1.5	Siemens
EP968	3700	108	0.9x0.9x7.25	256x256x20	5	1.5	Siemens
EP535	4813	48	1.8x1.8x4.5	128x128x27	4	3	Philips
EP774	8575	85	0.9x0.9x6	256x256x26	5	1.5	GE
EP815	4321	84	1.8x1.8x5	128x128x27	4	1.5	Philips
EP816	2604	68	0.79x0.79x5	288x288x27	4	3	Philips
EP829	3296	70	1.59x1.59x4	144x144x28	4	3	Philips
EP871	4616	100	1.6x1.6x4.4	144x144x30	4	1.5	Philips
EP902	2700	96	1.8x1.8x7	128x128x19	5	1.5	Siemens
EP936	3700	108	0.9x0.9x7	256x256x20	5	1.5	Siemens
EP988	4500	103	1.2x1.2x5.5	192x192x25	5	1.5	Siemens
EP1012	3600	107	0.9x0.9x7	256x256x20	5	1.5	Siemens
EP1111	2700	96	1.8x1.8x8	128x128x19	5	1.5	Siemens
EP1115	2700	96	1.8x1.8x6.5	128x128x19	5	1.5	Siemens
EP1117	2700	96	1.8x1.8x7.5	128x128x19	5	1.5	Siemens
MB532	2486	62	1.8x1.8x5	128x128x27	4	3	Philips
MB536	8000	88.6	0.94x0.94x6.5	256x256x20	5	1.5	GE
MB616	8000	94.6	0.98x0.98x4.4	256x256x30	4	1.5	GE
MB622	5826	58	2x2x3.3	112x112x32	3	3	Phillips
MB629	8575	74.6	0.9x0.9x6	256x256x30	5	1.5	GE
MB673	6200	161	1.8x1.8x6.5	128x128x22	5	1.5	Siemens
MB718	4411	84	1.8x1.8x5	128x128x27	4	1.5	Philips
MB719	4620	100	1.6x1.6x4.4	144x144x30	4	1.5	Philips
MB762	5335	83.6	1.8x1.8x5	128x128x32	4	1.5	Philips
MB783	3086	68	0.8x0.8x4.5	288x288x32	4	3	Philips

MB792	5545	89	1.2x1.2x4	192x192x35	4	1.5	Siemens
MB793	4542	100	1.6x1.6x4.4	144x144x30	4	1.5	Philips
MB796	8000	90	2x2x3.3	112x112x36	3	3	Philips
MB803	8575	84.6	0.9x0.9x5	256x256x33	5	1.5	GE
MB827	4452	89	1.8x1.8x4.4	128x128x30	4	1.5	Philips
MB834	5400	94	1.2x1.2x4	192x192x32	4	1.5	Siemens
MB844	7300	97	1.2x1.2x4	192x192x42	4	1.5	Siemens
MB860	8575	86	0.9x0.9x5	256x256x28	5	1.5	GE
MB861	8575	85	0.9x0.9x5	256x256x31	5	1.5	GE
MB862	6700	99	1.2x1.2x4	192x192x38	4	1.5	Siemens
MB929	6000	99	1.2x1.2x4	192x192x34	4	1.5	Siemens
MB1075	4500	103	1.2x1.2x5.5	192x192x25	5	1.5	Siemens
MB1098	8900	78	1.3x1.3x5	180x180x29	4	3	Siemens
PA571	3600	107	0.9x0.9x7.5	256x256x20	5	1.5	Siemens
PA593	5300	161	1.56x1.56x6.5	104x128x19	5	1.5	Siemens
PA607	8700	110	0.94x0.94x6.5	256x256x20	5	1.5	GE
PA613	4300	89	1.8x1.8x4.4	128x128x27	4	1.5	Philips
PA624	2605	68	0.79x0.79x5	288x288x27	4	3	Philips
PA645	5088	99	1.95x1.95x4.4	128x128x27	4	1.5	Philips
PA655	5026	89	1.8x1.8x4.4	128x128x31	4	1.5	Philips
PA666	5826	58	2x2x3.3	112x112x32	3	3	Philips
PA696	3176	57	2x2x3.3	112x112x32	3	3	Philips
PA722	3350	74	2.4x2.4x5	96x96x24	5	1.5	Philips
PA723	5826	58	2x2x3.3	112x112x32	3	3	Philips
PA728	8575	85	0.9x0.9x6	256x256x24	5	1.5	GE
PA731	3094	68	0.8x0.8x4.4	288x288x32	4	3	Philips
PA741	8575	85	0.9x0.9x6	256x256x23	5	1.5	GE
PA753	2990	62	1.6x1.6x5	144x144x32	4	3	Philips
PA784	2604	68	0.79x0.79x5	288x288x27	4	3	Philips
PA791	8000	85	0.9x0.9x5	256x256x29	5	1.5	GE
PA808	2703	68	0.9x0.9x5	256x256x28	4	3	Philips
PA832	5214	100	1.6x1.6x4.4	144x144x34	4	1.5	Philips
PA837	8000	92	0.98x0.98x4.4	256x256x29	4	1.5	GE
PA856	2903	68	0.9x0.9x5	256x256x30	4	3	Philips
PA863	7736	86	2.2x2.2x2.5	104x108x61	2.5	1.5	Siemens
PA873	7129	89	1.2x1.2x4	192x192x45	4	1.5	Siemens
PA883	3700	108	0.9x0.9x7	256x256x20	5	1.5	Siemens
PA903	2700	96	1.8x1.8x6.5	128x128x19	5	1.5	Siemens
PA912	2700	96	1.8x1.8x6.5	128x128x19	5	1.5	Siemens
PA934	6100	89	1.2x1.2x4	192x192x38	4	1.5	Siemens
PA954	3700	108	0.9x0.9x7.5	256x256x20	5	1.5	Siemens
PA961	3700	108	0.9x0.9x7.5	256x256x20	5	1.5	Siemens

Table C.4.2: Imaging parameters of T2 for experiment 4.

Case	TR	TE	Resolution	FOV	Slice thickness	B_0	Scanner
EP611	5050	100	0.45x0.45x5	512x512x25	4	1.5	Philips
EP639	3000	85	0.45x0.45x4.4	512x512x32	4	3	Philips
EP813	5300	84	0.43x0.43x6	512x512x23	5	1.5	GE
EP821	5100	85	0.45x0.45x4	512x512x46	3	1.5	GE
EP904	6180	115	0.66x0.66x6	240x320x22	4	1.5	Siemens
EP968	4000	89	0.43x0.43x7.25	384x512x20	5	1.5	Siemens
EP535	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
EP774	4220	85	0.45x0.45x4	512x512x38	3	1.5	GE
EP815	6055	100	0.4x0.4x4.4	512x512x30	4	1.5	Philips
EP816	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
EP829	3000	80	0.45x0.45x5	512x512x24	4	3	Philips
EP871	6055	100	0.4x0.4x4.4	512x512x30	4	1.5	Philips
EP902	4920	101	0.57x0.57x5.6	306x384x25	4	1.5	Siemens
EP936	4000	89	0.43x0.43x7.5	384x512x20	5	1.5	Siemens
EP988	4750	90	0.5x0.5x5.5	360x448x25	5	1.5	Siemens
EP1012	5550	30	0.4x0.4x8.5	384x512x16	5	1.5	Siemens
EP1111	4920	95	0.6x0.6x6.4	306x384x25	4	1.5	Siemens
EP1115	6180	14	0.63x0.63x5.8	240x320x22	4	1.5	Siemens
EP1117	6180	14	0.63x0.63x6.8	240x320x22	4	1.5	Siemens
MB532	3000	80	0.45x.45x5	512x512x27	4	3	Philips
MB536	5060	83.5	0.47x0.47x6.5	512x512x20	5	1.5	GE

MB616	3760	95.7	0.49x0.49x4.4	512x512x30	4	1.5	GE
MB622	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
MB629	5040	87.4	0.45x0.45x4	512x512x43	3	1.5	GE
MB673	4560	105	0.90x0.90x6.5	224x256x22	5	1.5	Siemens
MB718	6459	100	0.41x0.41x4.4	512x512x32	4	1.5	Philips
MB719	6070	100	0.45x0.45x4.4	512x512x30	4	1.5	Philips
MB762	6474	100	0.4x0.4x5	512x512x32	4	1.5	Philips
MB783	3000	85	0.45x0.45x4.5	512x512x32	4	3	Philips
MB792	6980	77	0.45x0.45x3.9	402x448x40	3	1.5	Siemens
MB793	6055	100	0.41x0.41x4.4	512x512x30	4	1.5	Philips
MB796	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
MB803	4660	85	0.45x0.45x4	512x512x41	3	1.5	GE
MB827	6059	100	0.45x0.45x4.4	512x512x30	4	1.5	Philips
MB834	5230	107	0.72x0.72x5.2	256x320x30	4	1.5	Siemens
MB844	6890	77	0.51x0.51x3.9	402x448x42	3	1.5	Siemens
MB860	4260	86	0.45x0.45x4	512x512x37	3	1.5	GE
MB861	4440	85	0.45x0.45x4	512x512x39	3	1.5	GE
MB862	6980	77	0.45x0.45x3.9	360x448x40	3	1.5	Siemens
MB929	6980	77	0.51x0.51x3.9	402x448x40	3	1.5	Siemens
MB1075	5130	90	0.51x0.51x5.5	360x448x27	5	1.5	Siemens
MB1098	5660	104	0.5x0.5x4.8	402x448x29	4	3	Siemens
PA571	4000	89	0.4x0.4x7.5	384x512x20	5	1.5	Siemens
PA593	4970	105	0.4x0.4x6.5	408x512x24	5	1.5	Siemens
PA607	5320	89	0.47x0.47x6.5	512x512x20	5	1.5	GE
PA613	5453	100	0.45x0.45x4.4	512x512x27	4	1.5	Philips
PA624	3000	80	0.44x0.44x5	480x480x27	4	3	Philips
PA645	5453	100	0.45x0.45x4.4	512x512x27	4	1.5	Philips
PA655	6261	100	0.45x0.45x4.4	512x512x31	4	1.5	Philips
PA666	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	3000	85	0.45x0.45x4.5	512x512x22	4	3	Philips
PA722	5595	110	0.45x0.45x6	512x512x25	5	1.5	Philips
PA723	3000	85	0.45x0.45x4.5	512x512x27	4	3	Philips
PA728	9140	86	0.45x0.45x4	512x512x41	3	1.5	GE
PA731	3000	80	0.45x0.45x4.4	512x512x34	4	3	Philips
PA741	2660	84	0.4x0.4x6	512x512x23	5	1.5	GE
PA753	3000	80	0.45x0.45x5	512x512x27	4	3	Philips
PA784	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA791	4440	85	0.45x0.45x4	512x512x39	3	1.5	GE
PA808	3000	85	0.45x0.45x4.5	512x512x30	4	3	Philips
PA832	6863	100	0.4x0.4x4.4	512x512x34	4	1.5	Philips
PA837	7120	95	0.49x0.49x4.4	512x512x29	4	1.5	GE
PA856	3000	85	0.45x0.45x5	512x512x30	4	3	Philips
PA863	6980	77	0.5x0.5x3.9	402x448x42	3	1.5	Siemens
PA873	7020	77	0.45x0.45x3.9	360x448x43	3	1.5	Siemens
PA883	5550	119	0.4x0.4x8.5	384x512x16	5	1.5	Siemens
PA903	4920	101	0.67x0.57x5.6	306x384x25	4	1.5	Siemens
PA912	4920	101	0.57x0.57x5.6	306x384x25	4	1.5	Siemens
PA934	8162	77	0.5x0.5x3.9	402x448x50	3	1.5	Siemens
PA954	4000	89	0.43x0.43x8	384x512x20	5	1.5	Siemens
PA961	4000	89	0.43x0.43x7.5	384x512x20	5	1.5	Siemens

Table C.4.3: Imaging parameters of post-contrast T1 for experiment 4.

Case	TR	TE	Resolution	FOV	Slice Thickness	B_0	Scanner
EP611	400	15	0.9x0.9x5	256x256x25	4	1.5	Philips
EP639	645	14	0.45x0.45x4.4	512x512x32	4	3	Philips
EP813	500	22	0.43x0.43x6	512x512x23	5	1.5	GE
EP821	500	21	0.45x0.45x6.5	512x512x28	5	1.5	GE
EP904	580	11	0.55x0.55x6.75	288x384x20	4.5	1.5	Siemens
EP968	697	14	0.43x0.43x7.5	384x512x20	5	1.5	Siemens
EP535	574	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP774	500	21	0.45x0.45x5	512x512x30	4	1.5	GE
EP815	591	10	0.9x0.9x4.4	256x256x32	4	1.5	Philips
EP816	565	14	0.45x0.45x4.5	512x512x27	4	3	Philips
EP829	8.2	3.77	0.84x0.84x1.85	288x288x60	0.8	3	Philips
EP871	300	10	0.9x0.9x4.4	256x256x30	4	1.5	Philips
EP902	597	15	0.58x0.58x6.6	400x400x22	6	1.5	Siemens

EP936	697	14	0.43x0.43x7.5	384x512x20	5	1.5	Siemens
EP988	512	8.7	0.7x0.7x5.5	256x320x25	5	1.5	Siemens
EP1012	697	14	0.4x0.4x7	384x512x20	5	1.5	Siemens
EP1111	580	11	0.6x0.6x7.9	288x384x20	4.5	1.5	Siemens
EP1115	580	11	0.55x0.55x6.75	288x384x20	4.5	1.5	Siemens
EP1117	580	11	0.55x0.55x7.65	288x384x16	4.5	1.5	Siemens
MB532	425	10	0.45x.45x5	512x512x27	4	3	Philips
MB536	660	22	0.47x0.47x6.5	512x512x20	5	1.5	GE
MB616	500	22	0.49x0.49x4.4	512x512x30	4	1.5	GE
MB622	571	14	0.45x0.45x4.5	512x512x27	4	3	Philips
MB629	560	21	0.45x0.45x6.5	512x512x28	5	1.5	GE
MB673	828	17	0.45x0.45x5.2	384x512x30	4	1.5	Siemens
MB718	325	10	0.9x0.9x4.4	256x256x32	4	1.5	Philips
MB719	676	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB762	594	10	0.9x0.9x5	256x256x32	4	1.5	Philips
MB783	651	14	0.45x0.45x4.5	512x512x32	4	3	Philips
MB792	414	17	0.78x0.78x5	256x208x30	4	1.5	Siemens
MB793	565	10	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB796	612	14	0.45x0.45x4.5	512x512x30	4	3	Philips
MB803	500	21	0.45x0.45x5.5	512x512x29	4	1.5	GE
MB827	675	12	0.9x0.9x4.4	256x256x30	4	1.5	Philips
MB834	414	17	0.72x0.72x5	260x320x30	4	1.5	Siemens
MB844	414	17	0.72x0.72x5	260x320x30	4	1.5	Siemens
MB860	400	21	0.45x0.45x6.5	512x512x23	5	1.5	GE
MB861	540	21	0.45x0.45x5	512x512x32	4	1.5	GE
MB862	414	17	0.72x0.72x5	260x320x30	3	1.5	Siemens
MB929	414	17	0.72x0.72x5	260x320x30	4	1.5	Siemens
MB1075	553	8.7	0.72x0.72x5.5	256x320x27	5	1.5	Siemens
MB1098	2640	9	0.72x0.72x4.8	256x320x29	4	3	Siemens
PA571	697	14	0.4x0.4x7.5	384x512x20	5	1.5	Siemens
PA593	663	17	0.45x0.45x6.5	384x512x24	5	1.5	Siemens
PA607	700	23	0.47x0.47x6.5	512x512x20	5	1.5	GE
PA613	603	12	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA624	567	14	0.44x0.44x5	480x480x27	4	3	Philips
PA645	603	12	0.9x0.9x4.4	256x256x27	4	1.5	Philips
PA655	692	12	0.9x0.9x4.4	256x256x31	4	1.5	Philips
PA666	603	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA696	250	2.3	0.45x0.45x5	512x512x30	4	3	Philips
PA722	656	15	0.9x0.9x6	256x256x20	5	1.5	Philips
PA723	566	14	0.45x0.45x4.5	512x512x27	4	3	Philips
PA728	560	21	0.45x0.45x5	512x512x33	4	1.5	GE
PA731	689	14	0.45x0.45x4.4	512x512x34	4	3	Philips
PA741	520	22	0.4x0.4x5	512x512x30	4	1.5	GE
PA753	565	14	0.45x0.45x5	512x512x27	4	3	Philips
PA784	604	14	0.45x0.45x4.5	512x512x30	4	3	Philips
PA791	540	21	0.45x0.45x5	512x512x32	4	1.5	GE
PA808	646	14	0.45x0.45x4.5	512x512x32	4	3	Philips
PA832	628	10	0.9x0.9x4.4	256x256x34	4	1.5	Philips
PA837	500	22	0.49x0.49x4.4	512x512x29	4	1.5	GE
PA856	530	10	0.45x0.45x5	512x512x30	4	3	Philips
PA863	470	17	0.7x0.7x5	260x320x33	4	1.5	Siemens
PA873	525	17	0.72x0.72x5	260x320x37	4	1.5	Siemens
PA883	697	14	0.4x0.4x7	384x512x20	5	1.5	Siemens
PA903	580	11	0.57x0.57x6.75	288x384x20	4.5	1.5	Siemens
PA912	697	14	0.39x0.39x7.5	384x512x20	5	1.5	Siemens
PA934	497	17	0.72x0.72x5.2	320x320x35	4	1.5	Siemens
PA954	697	14	0.43x0.43x7.5	384x512x20	5	1.5	Siemens
PA961	697	14	0.43x0.43x7.5	384x512x20	5	1.5	Siemens

APPENDIX D: ADDITIONAL EXAMPLE IMAGES

In this section, additional example images for results in section 8.4.4 are shown in the following figures.

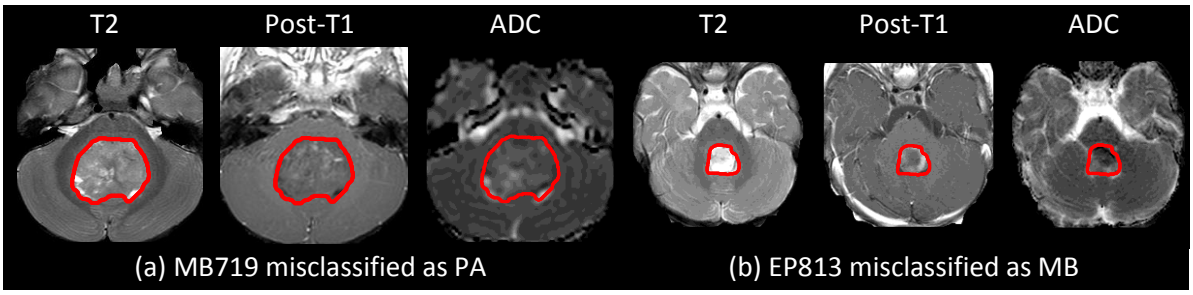


Figure D.1: Two cases misclassified by ADC based TA but correctly predicted by the superposition of weighted outcomes.

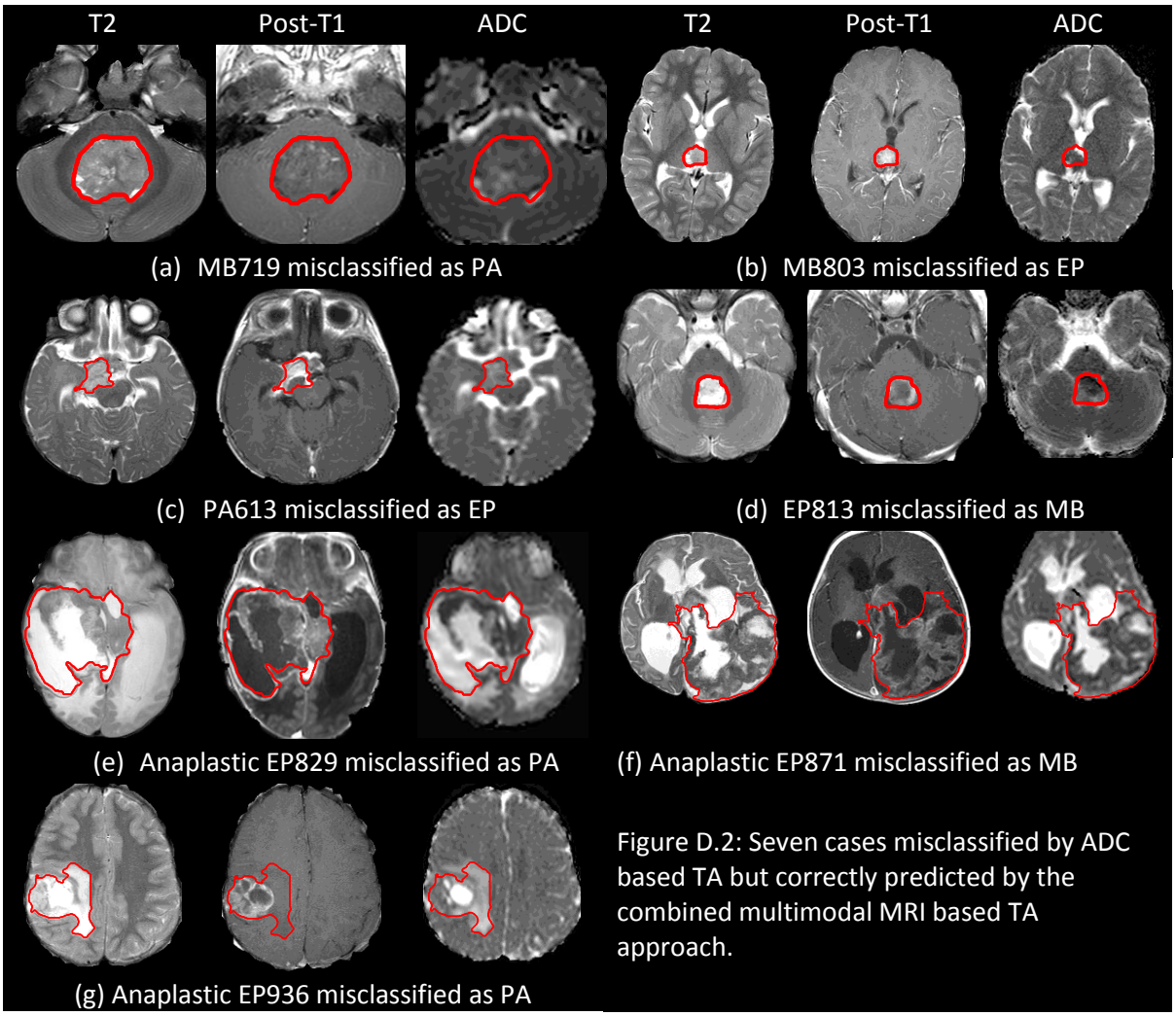


Figure D.2: Seven cases misclassified by ADC based TA but correctly predicted by the combined multimodal MRI based TA approach.

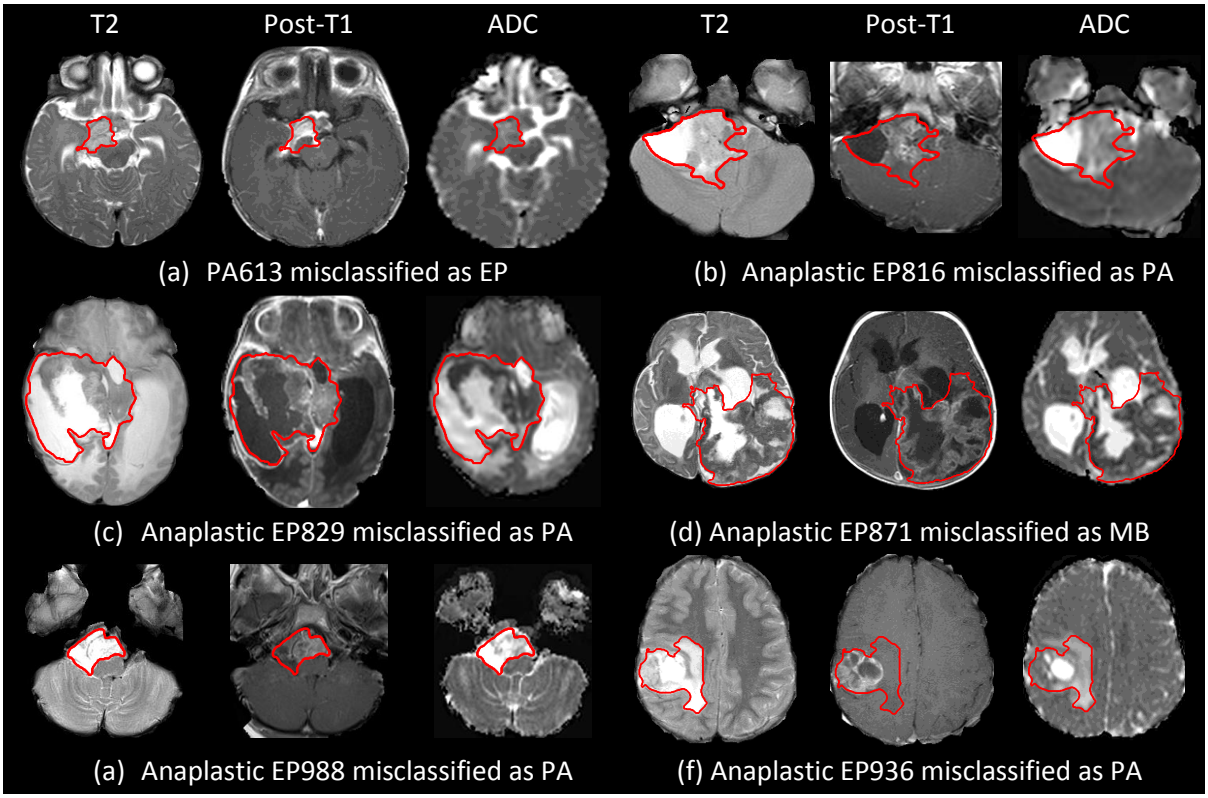


Figure D.3: Six cases misclassified by the superposition of weighted outcomes.

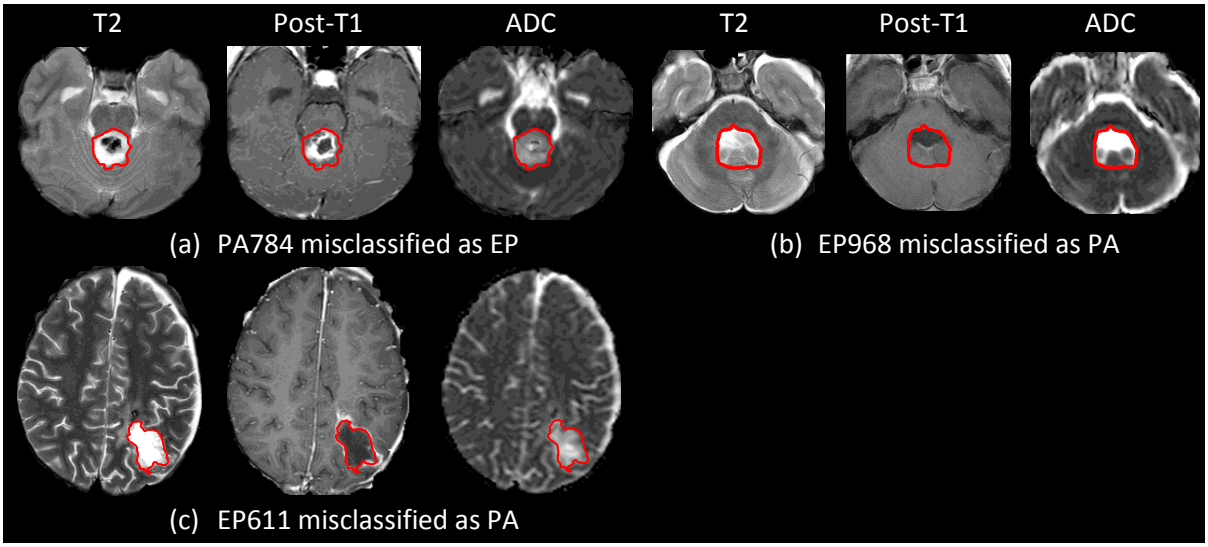


Figure D.4: Three cases misclassified by the combined multimodal MRI based TA approach.

APPENDIX E: PUBLICATION (BC ISMRM 2012)

Tantisatirapong, S., Davies, N.P., Peet, A.C., Arvanitis, T.N., 2012. Brain tumour segmentation based on multimodal magnetic resonance imaging in children, 21st British Chapter ISMRM Postgraduate Symposium, Wills Memorial Building, University of Bristol, Queens Road, Bristol BS8 1RJ.

Brain tumour segmentation based on multimodal magnetic resonance imaging in children

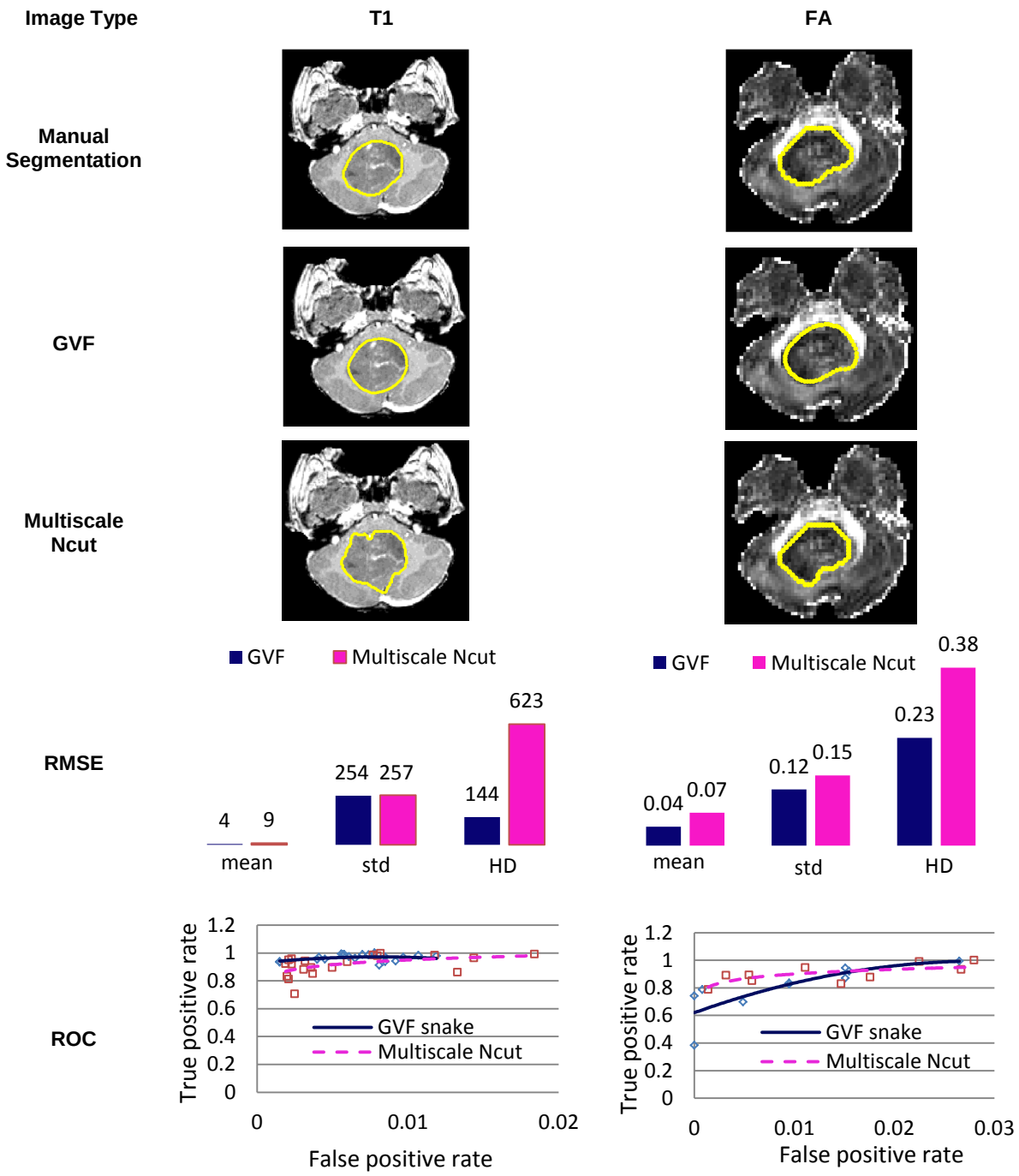
S Tantisatirapong[1], N P Davies[2,3,4] , A C Peet[2,3], T N Arvanitis[1,2]

[1] Electrical, Electronic and Computer Engineering, University of Birmingham, [2]Birmingham Children's Hospital, Birmingham, [3]Cancer Sciences, University of Birmingham, [4] Medical Physics and Imaging, University Hospital Birmingham

Introduction: Primary brain tumours are the most common solid tumours found in children and are an important cause of morbidity and mortality. MRI scanning is commonly used for non-invasive early-detection, diagnosis and delineation of tumours for treatment planning and assessment of post treatment changes. Different MRI modalities provide complementary contrast of tumour tissues with varying degrees of heterogeneity and diffusivity in different tumour types. Segmentation methods suitable for multimodal MRI are required to aid diagnosis, treatment planning and treatment response assessment. Hence, we aim to evaluate two well-known semi-automated segmentation methods named GVF snake and multiscale normalized cut and conduct the procedure based on conventional MRI images and Diffusion Tensor Imaging (DTI), which offers superior visualization of gross anatomical structure and tissue micro-structure respectively.

Methods: T1-weighted images (TE 3.37 ms, TR 1900 ms) of two patients were acquired using a 1.5T Siemens scanner. For each patient, fractional anisotropy (FA) maps were reconstructed using FSL tools [1] from a DTI dataset (TE 86 ms, TR 7000 ms, b-value 1000 s/mm^2, 42 gradient directions, resolution 2.22x2.22x2.50 mm).The T1-weighted and FA images were processed by the two methods. One is Gradient Vector Flow (GVF) deformable model or GVF snake [2] which is popularly used for medical image analysis. Initial seeds around tumour inserted by a user are a directive boundary. The method attempts to form edges by initially computing an edge map of an input image and progressing contour toward a force balance condition; balance an internal force (prevent stretching and bending of contour) and external force (pull the snake towards the desired contour) to form edges. The other is a multiscale normalized cut (ncut) method [3] which applies the normalized cut graph [4] as a partitioning framework. It computes the propagation of local cues (group of vertices formed by intensity similarity and intervening contours) across multiple ranges of spatial connections and offers the advantage of detecting coherent regions with a faint boundary as it appears in DTI images. The segmentation procedure was assessed by a process involving three levels of evaluation [5]. Firstly, a series of segmented images produced from the two methods were visually assessed. Secondly, a bootstrapping technique was used to optimize the selected methods from the first step. Finally, results obtained from the bootstrapping were compared with the gold standard (manual segments) by using empirical discrepancy methods. The root mean square error (RMSE) of mean and standard deviation of gray-level intensity within the region, Hausdorff distance (for shape matching) were computed from semi-automated and hand segments. It is expected that the closer the semi-automated segments to the gold standard, the lower the RMSE values. The receiver operating characteristics (ROC) was also used to characterize the performance of tests having lower RMSE values.

Results and Discussion: The T1-weighted and FA segments acquired from the two methods were visually monitored and shown in Figure 1 a)-d). The GVF snake could delineate the tumour when the initial seeds inserted close to the true boundary, while multiscale ncut performed segmenting according to the number of segment and internal parameter setting and produced reasonably results. Due to improved segmentation outcome and being more user-friendly, the multiscale ncut was selected for this study. The bootstrapping step based on multiscale ncut was performed by varying the main parameters of the method. ROC curves of test 1 and test 2, yielded lower RMSE values, were plotted in Figure 1 e)-f). Compare with test 2, the test 1, the blue ROC curves, provided higher true positive rate at the same false positive rate, and produced 99% accuracy for both T1 and FA based segmentation. Test 1 was selected for these data sets



Conclusions: The multiscale normalized cut offers the advantage of robustly segmenting low resolution and faint boundary of DTI based brain tumour images. It is a promising semi-automated technique to extract brain tumour for both FA and T1-weighted images close to the gold standard. However, the result is parameter-dependent and requires careful setting for satisfactory outcome.

Reference: [1] http://www.fmrib.ox.ac.uk/fsl/. [2] C Xu et al., (1997) IEEE Proc. Conf on Comp. Vis. Patt. Recog., 66-71. [3] T Cour et al., (2005) CVPR, 2:1124:31. [4] J Shi et al., (2000) IEEE Tran. Patt. Anal. Machine Intell.,22:888-905. [5] R. Pohle et al., (2002) Proc. SPIE 4684,287

APPENDIX F: PUBLICATION (BIOMED 2013)

Tantisatirapong, S., Davies N. P., Abernethy L., Auer D. P., Clark C. A., Grundy R., Jaspan T., Hargrave D., MacPherson L., Leach M. O., Payne G. S., Pizer B. L., Peet A. C., Arvanitis T. N., 2013a. Automated processing pipeline for texture analysis of childhood brain tumours based on multimodal magnetic resonance imaging, The IASTED International Conference on Biomedical Engineering (BioMed 2013). ACTA Press, Innsbruck, Austria, pp. 376-383.

AUTOMATED PROCESSING PIPELINE FOR TEXTURE ANALYSIS OF CHILDHOOD BRAIN TUMOURS BASED ON MULTIMODAL MAGNETIC RESONANCE IMAGING

Suchada Tantisatirapong[1], Nigel P. Davies[2,1,3], Lawrence Abernethy[4], Dorothee P. Auer[5,6], C. A. Clark[7], Richard Grundy[5,6], Tim Jaspan[6], Darren Hargrave[8], Lesley MacPherson[2], Martin O. Leach[8,9], Geoff S. Payne[8,9], Barry L. Pizer[4], Andrew C. Peet[1,3], Theodoros N. Arvanitis[1,3]

[1]University of Birmingham, Birmingham, United Kingdom, [2]University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom, [3]Birmingham Children's Hospital NHS Foundation Trust, Birmingham, United Kingdom, [4]Alder Hey Children's Hospital NHS Foundation Trust, Liverpool, United Kingdom, [5]University of Nottingham, Nottingham, United Kingdom, [6]University Hospital Nottingham, Nottingham, United Kingdom, [7]University College London, London, United Kingdom, [8]Royal Marsden Hospital, London, United Kingdom, [9]The Institute of Cancer Research, London, United Kingdom.

sxt096@bham.ac.uk[1],nigel.davies@nhs.net[2,1,3],Laurence.Abernethy@alderhey.nhs.uk[4],dorothee.auer@nottingham.ac.uk[5,6], c.clark@ich.ucl.ac.uk[7], Richard.Grundy@nottingham.ac.uk[5,6] , tim.jaspan@nuh.nhs.uk[6], darren.hargrave@rmh.nhs.uk[8], lesley.macpherson@bch.nhs.uk[2], martin@icr.ac.uk[8,9], gsp@icr.ac.uk[8,9], Barry.Pizer@alderhey.nhs.uk[4], a.peet@bham.ac.uk[1,3], t.arvanitis@bham.ac

ABSTRACT

Primary brain tumours are the most common solid tumours found in children and are an important cause of morbidity and mortality. Magnetic resonance imaging (MRI) is commonly used for non-invasive early-detection, diagnosis, delineation of tumours for treatment planning and assessment of post treatment changes. Different MRI modalities provide complementary contrast of tumour tissues, which can have varying degrees of heterogeneity and diffusivity in different tumour types. A variety of texture analysis methods have been shown to reveal tumour histological types. It is hypothesized that textural features, based on conventional and diffusion MRI modalities, would differentiate the characteristics of tumours. Tumour extraction is also a significant procedure needed to obtain a true tumour region. Semi-automated segmentation methods were applied, in comparison with the gold standard of manual segmentation by an expert, in order to speed up a manual segmentation approach and reduce any bias effects. In this study, we present an automatic processing pipeline for the characterization of brain tumours, based on texture analysis. We apply this to a multi-centre dataset of paediatric brain tumours and investigate the accuracy of

tumour classification, based on textural features of diffusion and conventional MR images.

KEY WORDS
Brain tumours, semi-automated segmentation, diffusion and conventional MRI, texture analysis, image processing

APPENDIX G: PUBLICATION (MEDICON 2013)

Tantisatirapong, S., Davies, N.P., Rodriguez, D., Abernethy, L., Auer, D.P., Clark, C.A., Grundy, R., Jaspan, T., Hargrave, D., MacPherson, L., Leach, M.O., Payne, G.S., Pizer, B.L., Peet, A.C., Arvanitis, T.N., 2013b. Magnetic resonance texture analysis: Optimal feature selection in classifying child brain tumors, XIII Mediterranean Conference on Medical and Biological Engineering and Computing. Springer, Seville, Spain, pp. 309-312.

Magnetic Resonance Texture Analysis: Optimal Feature Selection in Classifying Child Brain Tumors

Suchada Tantisatirapong[1], Nigel P. Davies[2,1,3], Daniel Rodriguez[4], Laurence Abernethy[5], Dorothee P. Auer[4,6], C. A. Clark[7,8], Richard Grundy[4,6], Tim Jaspan[6], Darren Hargrave[8], Lesley MacPherson[2], Martin O. Leach[9], Geoffrey S. Payne[9], Barry L. Pizer[5], Andrew C. Peet[1,3], Theodoros N. Arvanitis[1,3]

[1]University of Birmingham, Birmingham, United Kingdom, [2]University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom, [3]Birmingham Children's Hospital NHS Foundation Trust, Birmingham, United Kingdom, [4]University of Nottingham, Nottingham, United Kingdom, [5]Alder Hey Children's NHS Foundation Trust, Liverpool, United Kingdom, [6]University Hospital Nottingham, Nottingham, United Kingdom, [7]University College London, London, United Kingdom, [8]Great Ormond Street Hospital, London, United Kingdom, [9]The Institute of Cancer Research and Royal Marsden Hospital, Sutton, United Kingdom.

Abstract—**Textural feature based classification has shown that magnetic resonance images can characterize histological brain tumor types. Feature selection is an important process to acquire a robust textural feature subset and enhance classification rate. This work investigates two different feature selection techniques; principal component analysis (PCA), and the combination of max-relevance and min-redundancy (mRMR) and feedforward selection. We validated these techniques based on a multi-center dataset of pediatric brain tumor types; medulloblastoma, pilocytic astrocytoma and ependymoma, and investigated the accuracy of tumor classification, based on textural features of diffusion and conventional MR images.**

Keywords—**PCA, mRMR, feedforward selection, pediatric brain tumors, diffusion and conventional MR images**

I. INTRODUCTION

Conventional Magnetic Resonance Imaging (MRI) is an essential part of clinical diagnosis and offers superior structural imaging of anatomical elements compared with diffusion MRI. However, conventional MRI offers limited functional information of brain tissue. Diffusion MRI, a functional imaging technique, allows examination of tissue microstructure and provides complementary structural visualization. Diffusion weighted imaging (DWI) and diffusion tensor imaging (DTI), has been utilized as a biomarker in various neuro-pathological diseases [1]. A sensitization of DTI to subtle disturbances in white-matter tracts has revealed the neural tracts and tumor infiltration and invasion on healthy tissue [2]. Diffusion MR images would provide promising information for discriminating brain tumor types.

Texture analysis has been widely used to extract information from the MR images [3,4] Dimensionality reduction is commonly required to avoid over-fitting and to obtain the optimal feature subset. Choosing an appropriate feature selection approach is important because an inappropriate use of feature selection can distort classifier performance when larger datasets are used [5]. Feature selection is broadly divided into three main methods which are filter, wrapper and embedded approaches. Filter methods do not require learning algorithms; for example PCA, Fisher score and mutual information based feature selection. Wrapper and embedded methods use a learning algorithm to score feature subsets according to their discriminative rate as the selection criterion [6]. Wrapper methods can provide more accurate solutions than filter methods [7], but in general are more computationally expensive because the induction algorithm must be evaluated over each feature set. The embedded methods perform feature selection during the process of training. The most common embedded methods are such as least absolute shrinkage and selection operator (Lasso) and $\ell 1$-norm regularized Support Vector Machines [7]. In this study, we examine the unsupervised PCA and the combination of maximum relevance and minimum redundancy (mRMR) [8] and feedforward selection.

Support vector machines (SVMs), originally proposed by Vapnik et al 1992 [9], have shown superb performance at binary classification [10] and have been widely used in medical image analysis. SVMs tend to be more robust with smaller standard error compared to artificial neural network (ANN) training [11] and linear discriminant analysis (LDA), which is prone to non-singular value decomposition error. We applied the SVM as a validating classifier for the comparison of the two feature selection techniques.

II. MATERIAL AND MEDTHODS

A. Material

Three types of pre-treatment brain tumors have been considered: medulloblastoma (MB), pilocytic astrocytoma (PA) and ependymoma (EP). These cases were obtained from the Children's Cancer and Leukemia Group (CCLG) database. Apparent Diffusion Coefficient (ADC) maps reconstructed from DWI, Mean Diffusivity (MD) and Fractional Anisotropy (FA) maps derived from DTI and conventional MR images were acquired from four centers using a 1.5T GE, 1.5T Siemens, 1.5T and 3T Phillips scanners, following a common protocol defined by the CCLG Functional Imaging Group [12]. The number of MB, PA and EP cases used in this study is shown in Table 1.

Table 1 Number of subjects for each image type

Image type	MB	PA	EP	Resolution (x, y, z)
ADC (DWI)	16	21	11	(0.8, 0.8, 4) – (1.8, 1.8, 7.5)
MD (DTI)	19	17	8	(0.94, 0.94, 2.5) – (2, 2, 6.5)
FA (DTI)	19	17	8	(0.94, 0.94, 2.5) – (2, 2, 6.5)
T2w	25	34	15	(0.4, 0.4, 4) – (0.9, 0.9, 7.5)
FLAIR	20	19	9	(0.45, 0.45, 3) – (1, 1, 6.5)
T1w pre-contrast	21	24	12	(0.4, 0.4, 4.4) – (0.9, 0.9, 7.25)
T1w post-contrast	21	28	15	(0.4, 0.4, 4.4) – (0.9, 0.9, 7.25)

B. Data pre-processing

The processing pipeline is conducted according to the work presented by [13] including eddy current correction, skull stripping, Diffusion MR image reconstruction, registration, intensity normalization, segmentation and texture analysis. Due to protocol variation at the different centers and multiple slices acquired for the analysis, an image type would be included according to the slice spacing. If the registered images have the slice spacing greater than the T2w image (the reference image for registration and segmentation) and the difference is more than 20% by calculating from 100 x $(s_1-s_2)/(s_1+s_2)$, where s_1 is a T2w's slice spacing and s_2 is a current image's slice spacing, the images are excluded. The number of gray levels applied in the texture analysis affects the discrimination outcome. Inappropriate quantization provides unsatisfactory results and too many gray levels yields longer processing times. We set the number of gray levels for co-occurrence and run-length matrices to be 9 for all MR image types.

C. Feature selection

1. Principal Component Analysis (PCA)

PCA is an unsupervised feature selection method which transforms observed data to an orthogonal subspace such that the transformed variables are linearly uncorrelated (called principal components). The transformation arranges the principal components (PCs) in descending order in a way that the first PC has the largest variance; the second PC has the second largest variance and so on. In this study, we used the number of PCs that provides the maximum balanced accuracy (an average accuracy obtained from MB, PA and EP).

2. Maximum Relevance and Minimum Redundancy (mRMR)

The mRMR method applies mutual information theory in order to arrange features according to the criterion of maximum relevance (max-relevance) and minimum redundancy (min-redundancy). Max-relevance criteria search for features which satisfy the maximum relevant between individual features and class. The max relevant features could be highly redundant, because the features are highly dependent on each other. As a result, the power of discrimination would not change much. The highly redundant features are filtered by applying the min-redundancy condition to obtain the first m^{th} features. To obtain the highest accuracy, feedforward selection technique is used to search the optimal combination within the top-ranked 20th features. If any feature subset provides the same accuracy, the first combination feature subset is selected. For scoring feature in feedforwad selection, the accuracy is calculated by using balanced accuracy values.

D. Classification

Support vector machines (SVMs) compute the maximum marginal separation line between two classes. By using a kernel function, it maps the original features into higher dimensional space where it computes a hyperplane that maximizes the distance from the hyperplane to the training data in each class. Having found such a hyperplane, the SVM can then predict the classification of unlabeled test data by mapping it into the feature space and checking on which side of the separating plane the test data lies. We used the LIBSVM tools [14] with the function of C-SVC, and a linear kernel. For multi-classification, the 'one-against-one' approach implemented within LIBSVM was applied. Due to the small dataset, the validation of the classification was evaluated by using a leave-one-out validation technique.

III. RESULTS AND DISCUSSION

The number of principal components used is an important parameter in PCA analysis, with diffusion images requiring less components to achieve maximum classification rates than conventional images (see Fig. 1). This shows that the textural features derived from diffusion MR images contain the stronger feature subset in the top-ranked of PCs. In this study, features selected by using the PCA do not provide obviously better result than using the original feature set (see Fig. 2). The fact is that the PCA transforms all the observation data to obtain the maximum variance and still retains all variables to generate the PCs. Some outliers or noise, contributed by less separable textural features in the PCs, can affect the classification outcome.

Combining mRMR and feedforward selection scheme apparently enhances the discrimination performance for all of MR image types (see Fig. 2). The feature subset selected from this technique is similar to the finding presented by [15]. The group of most individually relevant features does not always provide the superior prediction result. Adding some less relevant features to the features having high discriminant index, contributes more to the accuracy of the classifier than using the top-ranked relevant features only (see Table 2).

The prediction rate for each tumor type (MB, PA and EP), as well as the overall accuracy, are higher in ADC and MD image based analysis (see Fig. 2). Textural features based on ADC and MD provides the most promising result compared to those based on other MR modalities. The result indicated a good separation between MB and PA. However improvement in the discrimination of EP from the other two groups still needs to be achieved.

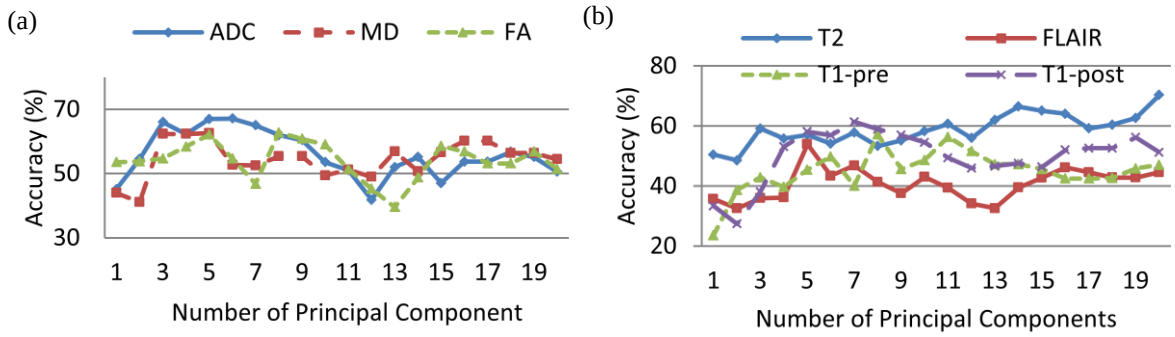


Fig. 1 Plot of number of principal component against classification accuracy for a) diffusion MR images. b) conventional MR images

Table 2 Feature ID selected from mRMR and feedforward selection for each MR image type

Image type	ID of the top-ranked 20th features selected by using mRMR	ID of the feature subset of the top 20th features selected by using feedforward selection
ADC (DWI)	[65, 62, 1, 97, 143, 121, 52, 85, 113, 49, 7, 9, 11, 123, 120, 94, 70, 42, 8, 57]	1-H:mean, 11-R:GLN 0°, 8-A:kurtosis, 9-R:SRE 0°, 65-C:svarh 0°, 57-C:cshad 0°, 42-R:SRE 135°, 85-C:svarh 45°
MD (DTI)	[1, 11, 80, 64, 56, 65, 5, 84, 2, 100, 104, 38, 12, 124, 129, 102, 6, 85, 24, 53]	1-H:mean, 56-C:cprom 0°, 12-R:RLN 0°
FA (DTI)	[142, 35, 81, 9, 134, 57, 70, 144, 28, 100, 44, 7, 21, 46, 91, 90, 19, 117, 139, 42]	134-Wavelet, 35-R:RP 90°, 28-R:SRHGE 45°, 142-Wavelet
T2w	[97, 86, 53, 139, 101, 77, 129, 10, 9, 117, 63, 126, 6, 22, 73, 8, 57, 131, 69, 83]	97-C:cshad 90°, 129-C:inf1h 135°, 117-C:cshad 135°, 57-C:cshad 0°, 139-Wavelet,8-A:kurtosis, 22-R:GLN 45°, 10-R:LRE 0°, 6-A:variance, 77-C:cshad 45°, 131-C:indnc 135°, 101-C:homom 90°, 73-C:autoc 45°, 69-C:inf1h 0°, 53-C:autoc 0°
FLAIR	[88, 44, 97, 94, 33, 27, 61, 129, 28, 3, 16, 7, 108, 6, 24, 17, 142, 38, 128, 74]	15-R:HGRE 0°, 97-C:cshad 90°, 44-R:GLN 135°, 3-H:skewness, 74-C:corrp 45°, 88-C:denth 45°, 95-C:contr 135°
T1w pre-contrast	[3, 60, 129, 2, 12, 145, 97, 102, 8, 56, 55, 17, 16, 57, 19, 135, 76, 1, 80, 23]	3-H:skewness, 102-C:maxpr 90°, 97-C:cshad 90°, 145-Wavelet, 76-C:cprom 45°, 129-C:inf1h 135°, 60-C:entropy 0°, 57-C:cshad 0°
T1w post-contrast	[57, 11, 96, 49, 143, 2, 47, 69, 22, 77, 30, 27, 100, 82, 56, 117, 144, 12, 17, 41]	56-C:cprom 0°, 22-R:GLN 45°, 47-R:LGRE 135°, 11-R:GLN 0°, 96-C:cprom 90°, 144-Wavelet, 30-R:LRHGE 45°, 77-C:cshad 45°, 49-R:SRLGE 135°, 57-C:cshad 0°, 41-R:LRHGE 90°, 117-C:cshad 135°

Abbreviation: H-histogram, A-absolute gradient, R-gray-level run length matrix, C-gray-level cooccurrence matrix, GLN-gray-level nonuniformity, SRE-short run emphasis, SRLGE-short run low gray-level emphasis, SRHGE-short run high gray-level emphasis, LRHGE-long run high gray-level emphasis, SRLGE-short run low gray-level emphasis, RLN-run length nonuniformity, LRE-long run empasis, HGRE-High Gray-Level Run Emphasis, RP-run percentage, svarh-sum variance, cshad-cluster shade, cprom-cluster prominence, inf1h-information measure of correlation, homom-homogeneity, indnc-inverse difference normalized, autoc-autocorrelation, denth-difference entropy, corrp-correlation, maxpr-maximum probability, contr-contrast [13],[16,17] .

IV. CONCLUSIONS

For all image types, combining mRMR and feedforward selection provides a textural feature subset that can better discriminate the three tumor types than using the original features or the transformed features extracted by using the PCA. With this feature selection approach, classification based on ADC and MD image analysis outperforms other types of MR images. Future work involves the investigation of fusing textural features, based on

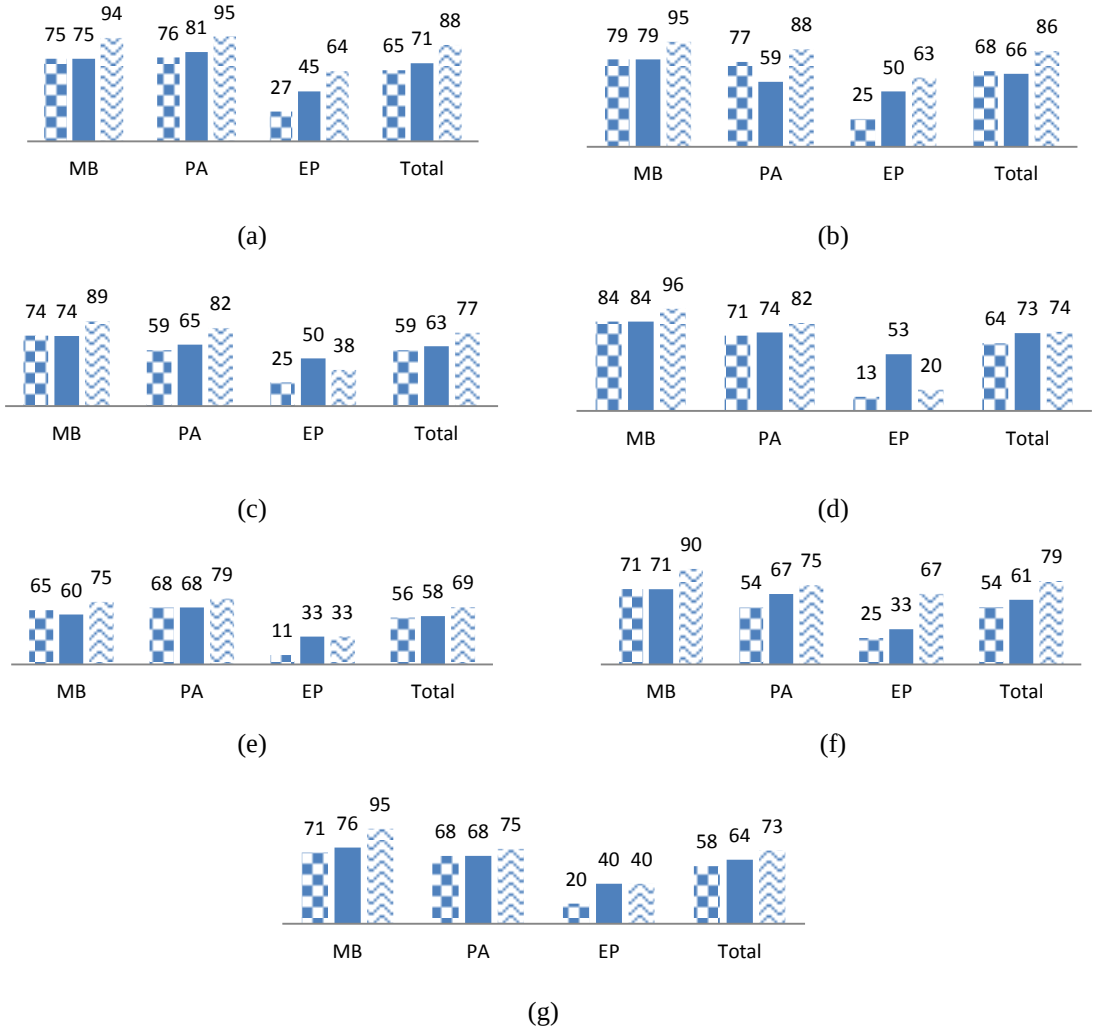


Fig. 2 The comparison of feature selection methods against the original feature based on (a) ADC, (b) MD, (c) FA, (d) T2, (e) FLAIR (f) T1w-pre contrast (g) T1w-post contrast. The legend: Original ■ PCA mRMR+Feedforward

multimodal MR images, in order to boost up the prediction performance of classifiers.

ACKNOWLEDGMENT

ST and TNA would like to acknowledge the financial support from the National Science and Technology Development Agency (NSTDA), Thailand. NPD is supported by a CSO/NIHR Healthcare Scientist Research Fellowship. DH, MOL, GSP acknowledge the support received from the CRUK and EPSRC Cancer Imaging Centre in association with the MRC and Department of Health (England) grant C1060/A10334, also NHS funding to the NIHR Biomedical Research Centre and NIHR Clinical Research Facility. MOL is an NIHR Senior Investigator. All authors acknowledge the support received from CRUK/EPSRC/MRC/NIHR Cancer Imaging Programme grant C7809/A10342.

REFERENCES

1. Oksuzler Y F, Cakmakci H, Kurul S et al. (2005) Diagnostic Value of Diffusion-Weighted Magnetic Resonance Imaging in Pediatric Cerebral Diseases. Pediatr Neurol 32:325-333
2. Price S, Jena, R, Burnet N et al. (2007) Predicting patterns of glioma recurrence using diffusion tensor imaging. Eur Radiol 17:1675-1684
3. Castellano G, Bonilha L, Li M et al. (2004) Texture analysis of medical images. Clin Radiol 59:1061-1069
4. Kassner A., and Thornhill, R. E. Texture analysis: a review of neurologic MR imaging applications. AJNR AM J Neuroradiol 31: 809-816
5. Smialowski P, Frishman D, Kramer S (2010) Pitfalls of supervised feature selection. Bioinformatics 26:440-443

6. Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. J Mach Learn Res 3:1157-1182
7. Peng H, Fulmi L, Ding C (2005) Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans. Pattern Anal. Mach. Intell. 27:1226-1238
8. Lijun Z, Chun C, Jiajun B et al. (2012) A Unified Feature and Instance Selection Framework Using Optimum Experimental Design. IEEE Trans. Image Process. 21:2379-2388
9. Boser B, Guyon L, Vapnik V (1992) A training algorithm for optimal margin classifiers. In the 5th Annual Workshop on Comp. Learning Theory, pp. 144-152, ACM, Pittsburgh, PA, USA, pp. 144-152
10. Li T, Zhu S, Ogihara M (2006) Using discriminant analysis for multi-class classification: an experimental investigation. Knowl. Inf. Syst. 10:453-472
11. Byvatov E, Fechner U, Sadowski J et al. (2003) Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. J. Chem. Inf. Comput. Sci. 43:1882-1889
12. http://www.cclgfig.bham.ac.uk/protocol.aspx
13. Tantisatirapong S, Davies N P, Abernethy L et al. (2013) Automated processing pipeline for texture analysis of childhood brain tumours based on multimodal magnetic resonance imaging. In the IASTED BioMed 2013, ACTA, Innsbruck, Austria, pp. 376-383,
14. Chang C-C, Lin C-J (2011) LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2:1-27
15. Krier C, François D, Wertz V et al. (2006) Feature Scoring by Mutual Information for Classification of Mass Spectra. In FLINS 2006, 7th Int. FLINS Conf. on Appl. AI, pp. 557-564
16. Xiaoou T (1998) Texture information in run-length matrices. IEEE Trans. Image Process. 7, 1602-1609
17. Haralick R M, Shanmugam K, Dinstein I H (1973) Textural Features for Image Classification. IEEE Trans. Syst. Man Cybern. Syst. 3, 610-621

APPENDIX H: PUBLICATION (ISMRM 2014)

Tantisatirapong, S., Davies, N.P., Rodriguez, D., Abernethy, L., Auer, D.P., Clark, C.A., Grundy, R., Jaspan, T., Hargrave, D., MacPherson, L., Leach, M.O., Payne, G.S., Pizer, B.L., Bailey, S., Peet, A.C., Arvanitis, T.N., 2014. Combining multi-centre conventional and diffusion MR texture for the characterisation of childhood brain tumours, Joint Annual Meeting ISMRM-ESMRMB 2014, Milan, Italy.

Combining Conventional and Diffusion MR Image Textural Features for the Characterization of Childhood Brain Tumors

Suchada Tantisatirapong[1], Nigel P. Davies[2,1,3], Daniel Rodriguez[4], Laurence Abernethy[5], Dorothee P. Auer[4,6], C. A. Clark[7,8], Richard Grundy[4,6], Tim Jaspan[6], Darren Hargrave[8], Lesley MacPherson[2], Martin O. Leach[9], Geoffrey S. Payne[9], Barry L. Pizer[5], Andrew C. Peet[1,3], Theodoros N. Arvanitis[10,3]

[1]University of Birmingham, Birmingham, UK, [2]University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK, [3]Birmingham Children's Hospital NHS Foundation Trust, Birmingham, UK, [4]University of Nottingham, Nottingham, UK, [5]Alder Hey Children's NHS Foundation Trust, Liverpool, UK, [6]University Hospital Nottingham, Nottingham, UK, [7]University College London, London, UK, [8]Great Ormond Street Hospital, London, UK, [9]The Institute of Cancer Research and Royal Marsden Hospital, Sutton, UK, [10]Institute of Digital Healthcare, WMG, University of Warwick, Coventry, UK

Abstract

Conventional magnetic resonance imaging (MRI) is an important tool in the diagnosis of childhood brain tumors. It has been commonly used to provide structural information of tumors; however it has shown a limited capacity to identify specific tumor types. Apparent Diffusion Coefficient (ADC), derived from MR-based Diffusion Weighted Imaging (DWI), provides complementary structural visualization and can also reveal the information relating to different histological types of brain tumors. Integrating multimodal MR image analysis can take advantage of such complementary information derived from both structural and diffusion MR images. In this paper, we present a supervised machine learning approach to achieve image analysis based on textural features from individual image types of T2, T1-post contrast (T1-post) and ADC, as well their combination, in order to characterize the most common pediatric brain tumors; medulloblastomas (MB), pilocytic astrocytomas (PA), and ependymomas (EP).

Material and Methods

A pre-treatment MR image dataset of pediatric brain tumors were considered, with the number of MB, PA, and EP being 23, 29, and 19 respectively. The acquisition of T2, T1-post and DWI was conducted in multiple centers using a 1.5T Siemens, 1.5T GE, 1.5T and 3T Phillips scanners. The processing pipeline followed the approach presented in [1] and the feature selection method followed the method suggested by [2]. A multinomial logistic regression technique [3] was customized as part of a classifier and validated by a leave-one-out cross validation. The texture based classification was performed on each single MR image type; T2, T1-post and ADC (reconstructed from DWI). Three approaches of integrating texture based T2, T1-post and ADC classifications were investigated. The first approach combines the textural features from all aforementioned image types. The second approach combines the probability outcome of the classifier obtained from each MR image type; the highest probability value obtained from the addition determines the final result. The third approach defines the integrated classification as a superposition of the weighted predicted class from each image type (i.e. this is the outcome with highest accuracy in the individual classifications of the ADC, T2 and T1-post image types).

Results and Discussion

The optimal textural feature set derived from ADC, T2, T1-post images individually and combined is shown in Table 1. In the single MR image type analysis, ADC based texture analysis (TA) can better discriminate the three tumor types compared to T2 and T1-post based texture analysis (see Figure 1). In the multimodal MR image analysis, the combining probability outcome tends to provide less accuracy. The

combined textural features and superposition of the weighted predicted class multimodal MR image analysis approaches give higher overall accuracy than the individual ADC based analysis. However, the superposition of the weighted predicted class approach shows, overall, a better separation between EP from MB and PA, while it still maintains a good classification for the MB and PA cases.

Table 1: The optimal textural feature set derived from single MR image and multimodal MR image based texture analysis

Image type	Textural Features
ADC	Csvarh(2,0^o), Csavgh(5,135^o), Hmean,Csosvh(4,0^o), Cidmnc(2,90^o), Ccshad(1,0^o)
T2	Csavgh(1,0^o), Wavelet Harr, Ccshad(1,90^o), Centro(1,90^o), Csvarh(1,90^o), Ccshad(2,90^o), Cdenth(2,135^o), Ccshad(5,45^o)
T1-post contrast	Ccprom(1,0^o), RLGRE 0 ^{o},RSRLGE 135°, Chomom (1,135^o)
Combined ADC, T2, and T1-post contrast	ADC: Csvarh(2,0^o), Csavgh(5,135^o), Hmean, Csosvh(4,0^o), RRP 45°, Cidmnc(2,90^o), RSRLGE 135°, T2: Centro(1,90), Wavelet Haar, Csvarh(1,90^o), Cdenth(2,135^o), Cinf1h(4,0 ^{o}), Csavgh(1,0 ^{o}), Ccshad(1,90^o), T1-post: Chomom(1,135^o), RSRLGE 135°, Wavelet Haar

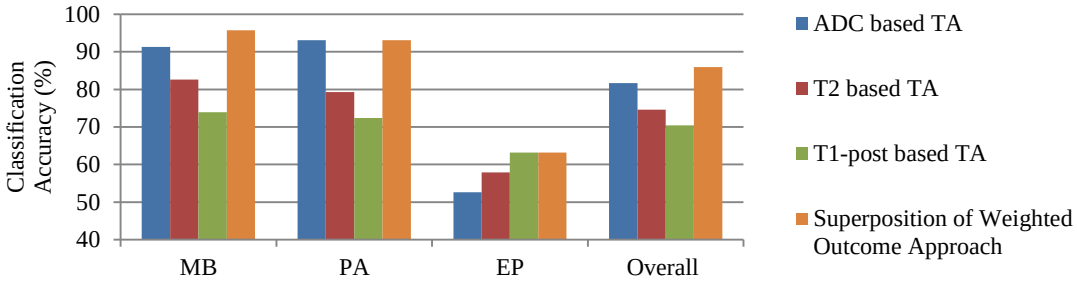


Figure 1: The comparison of classification accuracy obtained from single image type analysis and multiple image type analysis

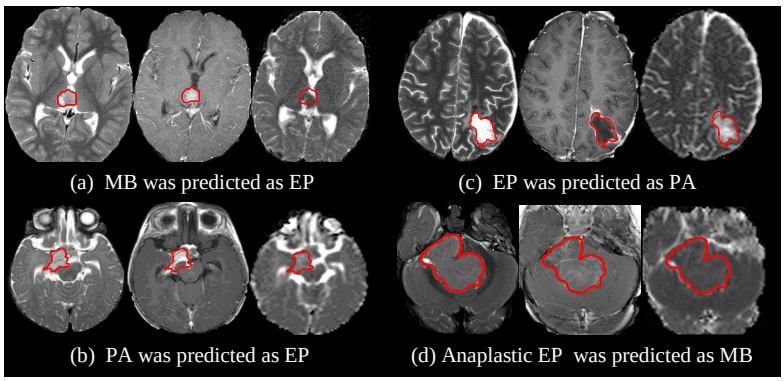


Figure 2: The example of four cases (out of ten cases) that were classified incorrectly by the superposition of weighted outcome approach. First, second and third image of each case is T2, T1-post and ADC image. The tumor boundary is delineated by red line.

Conclusion

In single image type analysis, ADC provided higher differentiation among the three brain tumor types. With this knowledge, the characterization of childhood brain tumor can be improved by analysing multimodal MR images based on the superposition of the weighted predicted class approach, with higher to lower weight given to ADC, T2 and T1-post respectively; the ependymomas can be better differentiate from the medulloblastomas and pilocytic astrocytomas.

Reference

[1] Tantisatirapong, S., et al., 2013. BioMed, 376-383 [2] Tantisatirapong, S., et al., 2013. MEDICON [3] McCullagh, P., et al., 1990. New York: Chapman & Hall

BIBLIOGRAPHY

1. Louis D, Ohgaki H, Wiestler O, Cavenee W, Burger P, Jouvet A, et al. The 2007 WHO Classification of Tumours of the Central Nervous System. Acta Neuropathologica. 2007;114(2):97-109.

2. Tecklin JS. Pediatric physical therapy. 4th ed: Lippincott Williams & Wilkins; 2008.

3. Saran F. Recent advances in paediatric neuro-oncology. Current Opinion in Neurology. 2002;15(6):671-7.

4. Cancer Research UK. CancerStats - Childhood Cancer - Great Britain & UK 2010 [cited 2014 Aug]. Available from: http://info.cancerresearchuk.org/cancerstats.

5. Becker LE. Pathology of pediatric brain tumors. Neuroimaging Clinics of North America. 1999;9(4):671-90.

6. Armstrong TS, Vera-Bolanos E, Bekele BN, Aldape K, Gilbert MR. Adult ependymal tumors: prognosis and the M. D. Anderson Cancer Center experience. Neuro-Oncology. 2010;12(8):862-70.

7. Johnson D, Brown P, Galanis E, Hammack J. Pilocytic astrocytoma survival in adults: analysis of the Surveillance, Epidemiology, and End Results Program of the National Cancer Institute. Journal of Neuro-Oncology. 2012;108(1):187-93.

8. Pan E, Prados MD. Adult Medulloblastomas: Holland-Frei Cancer Medicine. 6th edition. Hamilton (ON): BC Decker; 2003.

9. Panigrahy A, Bluml S. Neuroimaging of pediatric brain tumors: from basic to advanced magnetic resonance imaging (MRI). Journal of Child Neurology. 2009;24(11):1343-65.

10. Yuh EL, Barkovich AJ, Gupta N. Imaging of ependymomas: MRI and CT. Child's Nervous System. 2009;25(10):1203-13.

11. Le TH, Gean AD. Neuroimaging of traumatic brain injury. Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine. 2009;76(2):145-62.

12. Holland BA, Brant-Zawadzki M, Norman D, Hans Newton T. Magnetic resonance imaging of primary intracranial tumors: A review. International Journal of Radiation Oncology*Biology*Physics. 1985;11(2):315-21.

13. Galanis E, Buckner JC, Novotny P, Morton RF, McGinnis WL, Dinapoli R, et al. Efficacy of neuroradiological imaging, neurological examination, and symptom status in follow-up assessment of patients with high-grade gliomas. Journal of Neurosurgery. 2000;93(2):201-7.

14. Maier SE, Sun Y, Mulkern RV. Diffusion Imaging of Brain Tumors. NMR in biomedicine. 2010;23(7):849-64.

15. Sartor K. MR imaging of the brain: tumors. European Radiology. 1999;9(6):1047-54.

16. Brody AS, Frush DP, Huda W, Brent RL. Radiation risk to children from computed tomography. Pediatrics. 2007;120(3):677-82.

17. Arle JE, Morriss C, Wang ZJ, Zimmerman RA, Phillips PG, Sutton LN. Prediction of posterior fossa tumor type in children by means of magnetic resonance image properties, spectroscopy, and neural networks. Journal of Neurosurgery. 1997;86(5):755-61.

18. Kassner A, Thornhill RE. Texture Analysis: A Review of Neurologic MR Imaging Applications. AJNR American Journal of Neuroradiology. 2010;31(5):809-16.

19. Vlachou EO, Valchos N, Davies N, Arvanitis T, Grundy R, Peet A. Texture analysis of T1- and T2-weighted magnetic resonance images and use of probabilistic neural network to discriminate posterior fossa tumours in children. NMR in Biomedicine 2014.

20. Herlidou S, Rolland Y, Bansard JY, Le Rumeur E, de Certaines JD. Comparison of automated and visual texture analysis in MRI: Characterization of normal and diseased skeletal muscle. Magnetic Resonance Imaging. 1999;17(9):1393-7.

21. Nielsen B, Albregtsen F, Danielsen HE. Statistical nuclear texture analysis in cancer research: a review of methods and applications. Critical Reviews in Oncogenesis. 2008;14(2-3):89-164.

22. Zhang J, Tong L, Wang L, Li N. Texture analysis of multiple sclerosis: a comparative study. Magnetic Resonance Imaging. 2008;26:1160 - 6.

23. Holli K, Harrison L, Dastidar P, Waljas M, Liimatainen S, Luukkaala T, et al. Texture analysis of MR images of patients with Mild Traumatic Brain Injury. BMC Medical Imaging. 2010;10(1):8.

24. Herlidou-Même S, Constans JM, Carsin B, Olivie D, Eliat PA, Nadal-Desbarats L, et al. MRI texture analysis on texture test objects, normal brain and intracranial tumors. Magnetic Resonance Imaging. 2003;21(9):989-93.

25. Pavlisa G, Pavlisa G, Rados M. Diffusion differences between pilocytic astrocytomas and grade II ependymomas. Radiology and Oncology. 2011;45(2):97-101.

26. Kjaer L, Ring P, Thomsen C, Henriksen O. Texture analysis in quantitative MR imaging - tissue characterization of normal brain and intracranial tumours at 1.5T. Acta Radiologica. 1995;36:127 - 35.

27. Cha S. Neuroimaging in neuro-oncology. Neurotherapeutics. 2009;6(3):465-77.

28. Schneider JF, Viola A, Confort-Gouny S, Ayunts K, Le Fur Y, Viout P, et al. Infratentorial pediatric brain tumors: the value of new imaging modalities. Journal of Neuroradiology. 2007;34(1):49-58.

29. Castillo M, Mukherji SK. Diffusion-weighted imaging in the evaluation of intracranial lesions. Seminars in Ultrasound, CT and MRI. 2000;21(6):405-16.

30. Rodriguez Gutierrez D, Awwad A, Meijer L, Manita M, Jaspan T, Dineen RA, et al. Metrics and textural features of MRI diffusion to improve classification of pediatric posterior fossa tumors. AJNR American Journal of Neuroradiology. 2014;35:1-7.

31. Tantisatirapong S, Davies NP, Peet AC, Arvanitis TN, editors. Brain tumour segmentation based on multimodal magnetic resonance imaging in children. 21st British Chapter ISMRM Postgraduate Symposium; 2012; University of Bristol, Bristol.

32. Tantisatirapong S, Davies N. P., Abernethy L., Auer D. P., Clark C. A., Grundy R., et al. Automated processing pipeline for texture analysis of childhood brain tumours based on multimodal magnetic resonance imaging. 10th IASTED International Conference on Biomedical Engineering; Innsbruck, Austria: ACTA Press; 2013. p. 376-83.

33. Tantisatirapong S, Davies NP, Rodriguez D, Abernethy L, Auer DP, Clark CA, et al. Magnetic resonance texture analysis: Optimal feature selection in classifying child brain tumors. XIII Mediterranean Conference on Medical and Biological Engineering and Computing; Seville, Spain: Springer; 2013. p. 309-12.

34. Tantisatirapong S, Davies NP, Rodriguez D, Abernethy L, Auer DP, Clark CA, et al., editors. Combining multi-centre conventional and diffusion MR texture for the characterisation of childhood brain tumours. Joint Annual Meeting ISMRM-ESMRMB; 2014; Milan, Italy: ISMRM.

35. Purcell EM, Torrey HC, Pound RV. Resonance Absorption by Nuclear Magnetic Moments in a Solid. Physical Review. 1946;69(1-2):37-8.

36. Bloch F, Hansen WW, Packard M. The Nuclear Induction Experiment. Physical Review. 1946;70(7-8):474-85.

37. Nobelprize.org. The Nobel Prize in Physics 1952: Nobel Media AB Web. ; [cited 2014 Jan]. Available from: http://www.nobelprize.org/nobel_prizes/physics/laureates/1952/.

38. Nobelprize.org. The Nobel Prize in Chemistry 1991: Nobel Media AB Web. ; [cited 2014 Jan]. Available from: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1991/.

39. Cavendish M. Inventors and inventions: Paul Bernabeo; 2008.

40. Nobelprize.org. The Nobel Prize in Physiology or Medicine 2003: Nobel Media AB Web. ; [cited 2014 Jan]. Available from: http://www.nobelprize.org/nobel_prizes/medicine/laureates/2003/.

41. McRobbie DW, Moore EA, Graves MJ, Prince MR. MRI from picture to proton. 2nd ed: Cambridge: Cambridge University Press; 2007.

42. Santarelli MF. Basic physics of MR signal and image generation. In: Landini L, Positano V, Santarelli MF, editors. Advanced image processing in magnetic resonance imaging: CRC Press 2005.

43. Timothy E. J Behrens, Johansen-Berg H. Diffusion MRI: From quantitative measurement to in vivo neuroanatomy: ScienceDirect (Online service); 2009.

44. Mori S. Introduction to Diffusion Tensor Imaging Burlington: Elsevier Science B.V.; 2007.

45. Mori S, Zhang J. Principles of Diffusion Tensor Imaging and Its Applications to Basic Neuroscience Research. Neuron. 2006;51(5):527-39.

46. Liang Z-P, Lauterbur PC. Principles of magnetic resonance imaging: A signal processing perspective: Wiley-IEEE Press; 2000.

47. Hashemi RH, Bradley WG, Lisanti CJ. T1, T2 and T2*. MRI: The Basics: Lippincott Williams & Wilkins; 2010.

48. Hargreaves B. Rapid gradient-echo imaging. Journal of Magnetic Resonance Imaging. 2012;36(6):1300-13.

49. Jones J, Bashir U. Gradient echo sequences http://radiopaedia.org/articles/gradient-echo-sequences-1: Radiopedia.org; [cited 2014 Jan].

50. Rydberg JN, Riederer SJ, Rydberg CH, Jack CR. Contrast optimization of fluid-attenuated inversion recovery (FLAIR) imaging. Magnetic Resonance in Medicine. 1995;34(6):868-77.

51. Ortendahl DA, Hylton NM, Kaufman L, Crooks LE. Use of computers for increasing the information content of NMR imaging. 7th Annual Symposium on Computer Applications in Medical Care; 1983: IEEE; 1983. p. 787-90.

52. Hornak JP. The Basics of MRI: Interactive Learning Software; 1996 [cited 2014 Jan]. Available from: https://www.cis.rit.edu/htbooks/mri/.

53. Sodickson A. Chapter 18 Introductory MRI Physics. In: Newton HB, Jolesz FA, Malkin MG, Bourekas E, Christoforidis G, editors. Handbook of neuro-oncology neuroimaging: Academic Press; 2008. p. 128-35.

54. Fick A. V. On liquid diffusion. Philosophical Magazine Series 4. 1855;10(63):30-9.

55. Einstein A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Annalen der Physik. 1905;17:549-60.

56. Einstein A. Investigations on the Theory of Brownian Movement: New York: Dover; 1956.

57. Gass A, Niendorf T, Hirsch JG. Acute and chronic changes of the apparent diffusion coefficient in neurological disorders—biophysical mechanisms and possible underlying histopathology. Journal of the Neurological Sciences. 2001;186, Suppl 1(0):S15-S23.

58. Jones DK. Fundamentals of diffusion MR imaging. In: Gillard JH, D.Waldman A, Barker PB, editors. Clinical MR Neuroimaging: Diffusion, Perfusion and Spectroscopy: Cambridge University Press; 2005. p. 57.

59. Hahn EL. Spin Echoes. Physical Review. 1950;80(4):580-94.

60. Graessner J. Frequently asked questions: diffusion-weighted imaging (DWI): Siemens Healthcare, Hamburg, Germany; 2011 [cited Feb 2014]. Available from: www.siemens.com/magnetom-world.

61. Jones DK. Gaussian Modeling of Diffusion Signal. In: Johansen-Berg H, Behrens TEJ, editors. Diffusion MRI From quantitative measurement to in-vivo neuroanatomy Elsevier; 2009. p. 37-54.

62. Bito Y, Hirata S, Yamamoto E, editors. Optimal gradient factors for ADC measurements. 3rd Annual Meeting of ISMRM; 1995; France.

63. Chenevert TL, Brunberg JA, Pipe JG. Anisotropic diffusion in human white matter: demonstration with MR techniques in vivo. Radiology. 1990;177(2):401-5.

64. Doran M, Hajnal JV, Van Bruggen N, King MD, Young IR, Bydder GM. Normal and abnormal white matter tracts shown by MR imaging using directional diffusion weighted sequences. Journal of Computer Assisted Tomography. 1990;14(6):865-73.

65. Kingsley PB. Introduction to diffusion tensor imaging mathematics: Part I. Tensors, rotations, and eigenvectors. Concepts in Magnetic Resonance Part A. 2006;28A(2):101-22.

66. Price S, Jena R, Burnet N, Carpenter T, Pickard J, Gillard J. Predicting patterns of glioma recurrence using diffusion tensor imaging. European Radiology. 2007;17(7):1675-84.

67. Anderson AW. Theoretical analysis of the effects of noise on diffusion tensor imaging. Magnetic Resonance in Medicine. 2001;46(6):1174-88.

68. Leemans A, Jones DK. The B-matrix must be rotated when correcting for subject motion in DTI data. Magnetic Resonance in Medicine. 2009;61(6):1336-49.

69. Beaulieu C. The basis of anisotropic water diffusion in the nervous system – a technical review. NMR in Biomedicine. 2002;15(7-8):435-55.

70. Assaf Y, Pasternak O. Diffusion tensor imaging (DTI)-based white matter mapping in brain research: a review. Journal of Molecular Neuroscience. 2008;34(1):51-61.

71. Dougherty MJ, Santi M, Brose MS, Ma C, Resnick AC, Sievert AJ, et al. Activating mutations in BRAF characterize a spectrum of pediatric low-grade gliomas. Neuro-Oncology. 2010;12(7):621-30.

72. de Carvalho Neto A, Gasparetto EL, Ono SE, Bertoldi GA, Gomes AF. Adult cerebellar medulloblastoma: CT and MRI findings in eight cases. Arq Neuropsiquiatr. 2003;61(2A):199-203.

73. Godfraind C. Classification and controversies in pathology of ependymomas. Child's Nervous System. 2009;25(10):1185-93.

74. Eran A, Ozturk A, Aygun N, Izbudak I. Medulloblastoma: atypical CT and MRI findings in children. Pediatric Radiology. 2010;40(7):1254-62.

75. Gibson P, Tong Y, Robinson G, Thompson MC, Currle DS, Eden C, et al. Subtypes of medulloblastoma have distinct developmental origins. Nature. 2010;468(7327):1095-9.

76. Crawford JR, MacDonald TJ, Packer RJ. Medulloblastoma in childhood: new biological advances. The Lancet Neurology. 2007;6(12):1073-85.

77. Collins VP. Brain tumours: classification and genes. Journal of Neurology, Neurosurgery & Psychiatry. 2004;75(suppl 2):ii2-ii11.

78. Liu H-Q, Yin X, Li Y, Zhang J, Wang Y, Tchoyoson Lim CC, et al. MRI features in children with desmoplastic medulloblastoma. Journal of Clinical Neuroscience. 2012;19(2):281-5.

79. Martínez León MI. Review and update about medulloblastoma in children. Radiología (English Edition). 2011;53(2):134-45.

80. Meyers SP, Kemp SS, Tarr RW. MR imaging features of medulloblastomas. American Journal of Roentgenology. 1992;158(4):859-65.

81. Bühring U, Strayle-Batra M, Freudenstein D, Scheel-Walter HG, Küker W. MRI features of primary, secondary and metastatic medulloblastoma. European Radiology. 2002;12(6):1342-8.

82. Arai K, Sato N, Aoki J, Yagi A, Taketomi-Takahashi A, Morita H, et al. MR signal of the solid portion of pilocytic astrocytoma on T2-weighted images:is it useful for differentiation from medulloblastoma? Neuroradiology. 2006;48(4):233-7.

83. Chawla A, Emmanuel JV, Seow WT, Lou J, Teo HE, Lim CCT. Paediatric PNET: pre-surgical MRI features. Clinical Radiology. 2007;62(1):43-52.

84. Forbes JA, Reig AS, Smith JG, Jermakowicz W, Tomycz L, Shay SD, et al. Findings on preoperative brain MRI predict histopathology in children with cerebellar neoplasms. Pediatric Neurosurgery. 2011;47(1):51-9.

85. Lee YY, Tassel PVT, Bruner JM, Moser RP, Share JC. Juvenile pilocytic astrocytomas: CT and MR characteristics. American Journal of Roentgenology. 1989;152(6):1263-70.

86. Sharma MK, Mansur DB, Reifenberger G, Perry A, Leonard JR, Aldape KD, et al. Distinct genetic signatures among pilocytic astrocytomas relate to their brain region origin. Cancer Research. 2007;67(3):890-900.

87. Mahajan M, Sharma R, Sharma P, Gupta A. Densely Calcified Pilocytic Astrocytoma in the Sellar/Suprasellar Region. International Journal of Clinical Pediatrics. 2012;1(4-5):129-32.

88. Bükte Y, Paksoy Y, Genç E, Uca AU. Role of diffusion-weighted MR in differential diagnosis of intracranial cystic lesions. Clinical Radiology. 2005;60(3):375-83.

89. Missler U, Zimmermann M, Grehl H. Atypical pilocytic astrocytoma of the third ventricle appearing as a colloid cyst. Clinical Neuroradiology. 2009;19(4):301-5.

90. Choi J-Y, Chang K-H, Yu IK, Kim KH, Kwon BJ, Han MH, et al. Intracranial and spinal ependymomas: Review of MR images in 61 patients. Korean Journal of Radiology. 2002;3(4):219-28.

91. Martínez León MI, Vidal Denis M, Weil Lara B. Magnetic resonance imaging of infratentorial anaplastic ependymoma in children. Radiología (English Edition). 2012;54(1):59-64.

92. Vinchon M, Soto-Ares G, Riffaud L, Ruchoux MM, Dhellemmes P. Supratentorial ependymoma in children. Pediatric Neurosurgery. 2001;34(2):77-87.

93. Smith AB, Smirniotopoulos JG, Horkanyne-Szakaly I. From the radiologic pathology archives: Intraventricular neoplasms: Radiologic-pathologic correlation. RadioGraphics. 2013;33(1):21-43.

94. Tsuchiya K, Mizutani Y, Hachiya J. Preliminary evaluation of fluid-attenuated inversion-recovery MR in the diagnosis of intracranial tumors. AJNR American Journal of Neuroradiology. 1996;17(6):1081-6.

95. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. Medical Image Analysis. 2001;5(2):143-56.

96. Jezzard P, Balaban RS. Correction for geometric distortion in echo planar images from B0 field variations. Magnetic Resonance in Medicine. 1995;34(1):65-73.

97. Alexander AL, Lee JE, Lazar M, Field AS. Diffusion tensor imaging of the brain. Neurotherapeutics. 2007;4(3):316-29.

98. Metaxas D, Axel L, Fichtinger G, Székely G, Wu M, Chang L-C, et al. Comparison of EPI distortion correction methods in diffusion tensor MRI using a novel framework. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. Lecture Notes in Computer Science. 5242: Springer Berlin / Heidelberg; 2008. p. 321-9.

99. Poot DH, Pintjens W, Verhoye M, Van Der Linden A, Sijbers J. Improved B(0) field map estimation for high field EPI. Magnetic Resonance Imaging. 2010;28(3):441-50.

100. Funai AK, Fessler JA, Yeo D, Olafsson VT, Noll DC. Regularized field map estimation in MRI. IEEE Transactions on Medical Imaging. 2008;27(10):1484-94.

101. Le Bihan D, Poupon C, Amadon A, Lethimonnier F. Artifacts and pitfalls in diffusion MRI. Journal of Magnetic Resonance Imaging. 2006;24(3):478-88.

102. Jezzard P, Barnett AS, Pierpaoli C. Characterization of and correction for eddy current artifacts in echo planar diffusion imaging. Magnetic Resonance in Medicine. 1998;39(5):801-12.

103. Bastin ME. Correction of eddy current-induced artefacts in diffusion tensor imaging using iterative cross-correlation. Magnetic Resonance Imaging. 1999;17(7):1011-24.

104. Wan X, Gullberg GT, Parker DL, Zeng GL. Reduction of geometric and intensity distortions in echo-planar imaging using a multireference scan. Magnetic Resonance in Medicine. 1997;37(6):932-42.

105. Alexander AL, Tsuruda JS, Parker DL. Elimination of eddy current artifacts in diffusion-weighted echo-planar images: the use of bipolar gradients. Magnetic Resonance in Medicine. 1997;38(6):1016-21.

106. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. NeuroImage. 1999;9(2):179-94.

107. Sadananthan SA, Zheng W, Chee MW, Zagorodnov V. Skull stripping using graph cuts. NeuroImage. 2010;49(1):225-39.

108. Leung KK, Barnes J, Modat M, Ridgway GR, Bartlett JW, Fox NC, et al. Brain MAPS: An automated, accurate and robust brain extraction technique using a template library. NeuroImage. 2011;55(3):1091-108.

109. Clarke LP, Velthuizen RP, Camacho MA, Heine JJ, Vaidyanathan M, Hall LO, et al. MRI segmentation: Methods and applications. Magnetic Resonance Imaging. 1995;13(3):343-68.

110. Shattuck DW, Sandor-Leahy SR, Schaper KA, Rottenberg DA, Leahy RM. Magnetic resonance image tissue classification using a partial volume model. NeuroImage. 2001;13(5):856-76.

111. Smith SM. Fast robust automated brain extraction. Human Brain Mapping. 2002;17(3):143-55.

112. Segonne F, Dale AM, Busa E, Glessner M, Salat D, Hahn HK, et al. A hybrid approach to the skull stripping problem in MRI. NeuroImage. 2004;22(3):1060-75.

113. Boesen K, Rehm K, Schaper K, Stoltzner S, Woods R, Luders E, et al. Quantitative comparison of four brain extraction algorithms. NeuroImage. 2004;22(3):1255-61.

114. Albert H, Rafeef A, Roger T, Anthony T. MRI brain extraction with combined expectation maximization and geodesic active contours. IEEE International Symposium on Signal Processing and Information Technology; Aug. 2006; Vancouver, BC: IEEE; 2006. p. 107-11.

115. Zitová B, Flusser J. Image registration methods: a survey. Image and Vision Computing. 2003;21(11):977-1000.

116. Crum WR, Hartkens T, Hill DL. Non-rigid image registration: theory and practice. The British Journal of Radiology. 2004;77 Spec No 2:S140-53.

117. Jenkinson M. Image registration and motion correction. 2006.

118. Hill DL, Batchelor PG, Holden M, Hawkes DJ. Medical image registration. Physics in Medicine and Biology. 2001;46(3):R1-45.

119. Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G. Diffusion tensor MR imaging of the human brain. Radiology. 1996;201(3):637-48.

120. Jenkinson M, Bannister P, Brady M, Smith S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. NeuroImage. 2002;17(2):825-41.

121. Woods RP, Mazziotta JC, Cherry SR. MRI-PET registration with automated algorithm. Journal of Computer Assisted Tomography. 1993;17(4):536-46.

122. Studholme C, Hill DLG, Hawkes DJ. Automated 3-D registration of MR and CT images of the head. Medical Image Analysis. 1996;1(2):163-75.

123. Roche A, Malandain G, Pennec X, Ayache N. The correlation ratio as a new similarity measure for multimodal image registration. In: Wells W, Colchester A, Delp S, editors. Medical Image Computing and Computer-Assisted Interventation. Lecture Notes in Computer Science. 1496: Springer Berlin Heidelberg; 1998. p. 1115-24.

124. Shah M, Xiao Y, Subbanna N, Francis S, Arnold DL, Collins DL, et al. Evaluating intensity normalization on MRIs of human brain with multiple sclerosis. Medical Image Analysis. 2011;15(2):267-82.

125. Belaroussi B, Milles J, Carme S, Zhu YM, Benoit-Cattin H. Intensity non-uniformity correction in MRI: Existing methods and their validation. Medical Image Analysis. 2006;10(2):234-46.

126. Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. Journal of Magnetic Resonance Imaging. 2003;22:81 - 91.

127. Cox IJ, Roy S, Hingorani SL, editors. Dynamic histogram warping of image pairs for constant image brightness. International Conference on Image Processing; 1995 23-26 Oct. 1995.

128. Bergeest J-P, Jäger F. A comparison of five methods for signal intensity standardization in MRI. In: Tolxdorff T, Braun J, Deserno T, Horsch A, Handels H, Meinzer H-P, editors. Bildverarbeitung für die Medizin 2008. Informatik aktuell: Springer Berlin Heidelberg; 2008. p. 36-40.

129. Wells WM, III, Grimson WEL, Kikinis R, Jolesz FA. Adaptive segmentation of MRI data. IEEE Transactions on Medical Imaging. 1996;15(4):429-42.

130. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Transactions on Medical Imaging. 1998;17(1):87-97.

131. Nyul LG, Udupa JK. On standardizing the MR image intensity scale. Magnetic Resonance in Medicine. 1999;42(6):1072-81.

132. Hellier P. Consistent intensity correction of MR images. International Conference on Image Processing; 14-17 Sept. 2003: IEEE; 2003. p. I-1109-12.

133. Christensen JD. Normalization of brain magnetic resonance images using histogram even-order derivative analysis. Magnetic Resonance Imaging. 2003;21(7):817-20.

134. Weisenfeld NI, Warfteld SK. Normalization of joint image-intensity statistics in MRI using the Kullback-Leibler divergence. IEEE International Symposium on Biomedical Imaging: Nano to Macro; 15-18 April 2004: IEEE; 2004. p. 101-4.

135. Kullback S, Leibler RA. On information and sufficiency. The Annals of Mathematical Statistics. 1951;2(1):79-86.

136. Jäger F, Deuerling-Zheng Y, Frericks B, Wacker F, Hornegger J. A new method for MRI intensity standardization with application to lesion detection in the brain. Vision Modeling and Visualization 2006:269–76.

137. Withey DJ, Koles ZJ. Medical image segmentation: Methods and software. Joint Meeting of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging; 12-14 Oct. 2007: IEEE; 2007. p. 140-3.

138. Donnell OL. Semi-automatic medical image segmentation: Massachusetts institute of technology 2001.

139. Yu Jin Z. A review of recent evaluation methods for image segmentation. 6th International Symposium on Signal Processing and its Applications; 13-16 Aug 2001; Kuala Lumpur IEEE; 2001. p. 148-51.

140. McInerney T, Terzopoulos D. Deformable models. In: Bankman IN, editor. Handbook of medical image processing and analysis. 2nd ed: Elsevier; 2008.

141. Dawant BM, Zijdenbos AP. Image segmentation In: Sonka M, Fitzpatrick JM, editors. Handbook of medical imaging, Volume 2 - Medical image processing and analysis: SPIE Press; 2009.

142. Ma Z, Tavares JMRS, Jorge RMN, editors. A review on the current segmentation algorithms for medical images. 1st International Conference on Computer Imaging Theory and Applications; 2009; Lisboa, Portugal.

143. Patil DD, Deore SG. Medical image segmentation: A review. International Journal of Computer Science and Mobile Computing. 2013;2(1):22-7.

144. Canny J. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1986;PAMI-8(6):679-98.

145. Sobel I, Feldman G. A 3x3 isotropic gradient operator for image processing. 1968.

146. Bock H-H. Clustering methods: A history of k-Means algorithms. In: Brito P, Cucumel G, Bertrand P, Carvalho F, editors. Selected Contributions in Data Analysis and Classification. Studies in Classification, Data Analysis, and Knowledge Organization: Springer Berlin Heidelberg; 2007. p. 161-72.

147. Jayadevappa D, Srinivas Kumar S, Murty DS. A hybrid segmentation model based on watershed and gradient vector flow for the detection of brain tumor. International Journal of Signal Processing, Image Processing and Pattern Recognition. 2009;2(3):29-42.

148. Meyer F. Topographic distance and watershed lines. Signal Processing. 1994;38(1):113-25.

149. Rogowska J. Overview and fundamentals of medical image segmentation. In: Bankman IN, editor. Handbook of medical image processing and analysis. Second ed: Elsevier; 2008.

150. Dagher I, Tom KE. WaterBalloons: A hybrid watershed Balloon Snake segmentation. Image and Vision Computing. 2008;26(7):905-12.

151. Ma Z, Tavares JM, Jorge RN, Mascarenhas T. A review of algorithms for medical image segmentation and their applications to the female pelvic cavity. Computer Methods Biomechics and Biomedical Engineering. 2010;13(2):235-46.

152. Kazerooni AF, Ahmadian A, Serej ND, Rad HS, Saberi H, Yousefi H, et al. Segmentation of brain tumors in MRI images using multi-scale gradient vector flow. 33rd Annual International IEEE Engineering in Medicine and Biology Society; 30 Aug. 2011-3 Sept. 2011 Boston, MA IEEE; 2011. p. 7973-6.

153. Wang G, Wang D. Segmentation of brain MRI image with GVF snake model. 1st International Conference on Pervasive Computing Signal Processing and Applications 17-19 Sept. 2010; Harbin: IEEE; 2010. p. 711-4.

154. Chenyang X, Prince JL. Gradient vector flow: a new external force for snakes. IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 17-19 Jun. 1997; San Juan IEEE; 1997. p. 66-71.

155. Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. International Journal of Computer Vision. 1988;1(4):321-31.

156. Chenyang X, Prince JL. Snakes, shapes, and gradient vector flow. IEEE Transactions on Image Processing. 1998;7(3):359-69.

157. Zhang M, Li Q, Li L, Bai P. An improved algorithm based on the GVF-snake for effective concavity edge detection. Journal of Software Engineering and Applications. 2013;6(4):174-8.

158. Cour T, Benezit F, Shi J. Spectral segmentation with multiscale graph decomposition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 20-25 Jun. 2005: IEEE; 2005. p. 1124-31.

159. Shi J, Malik J. Normalized Cuts and Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000;22(8):888-905.

160. Courant R, Hilbert D. Methods of mathematical physics: John Wiley & Sons, Inc.; 1989.

161. Zhang H, Fritts JE, Goldman SA. Image segmentation evaluation: A survey of unsupervised methods. Computer Vision and Image Understanding. 2008;110(2):260-80.

162. Xie K, Yang J, Zhang ZG, Zhu YM. Semi-automated brain tumor and edema segmentation using MRI. European Journal of Radiology. 2005;56(1):12-9.

163. Archip N, Jolesz FA, Warfield SK. A validation framework for brain tumor segmentation. Academic Radiology. 2007;14(10):1242-51.

164. Brown R, Zlatescu M, Sijben A, Roldan G, Easaw J, Forsyth P, et al. The use of magnetic resonance imaging to noninvasively detect genetic signatures in oligodendroglioma. Clinical Cancer Research. 2008;14(8):2357-62.

165. Artzi M, Aizenstein O, Jonas-Kimchi T, Myers V, Hallevi H, Ben Bashat D. FLAIR lesion segmentation: Application in patients with brain tumors and acute ischemic stroke. European Journal of Radiology. 2013;82(9):1512-8.

166. Prastawa M, Bullitt E, Ho S, Gerig G. A brain tumor segmentation framework based on outlier detection. Medical Image Analysis. 2004;8(3):275-83.

167. McMillan KM, Ehtesham M, Stevenson CB, Edgeworth ML, Thompson RC, Price RR. T2 detection of tumor invasion within segmented components of glioblastoma multiforme. Journal of Magnetic Resonance Imaging. 2009;29(2):251-7.

168. Ahmed S, Iftekharuddin KM, Vossough A. Efficacy of texture, shape, and intensity feature fusion for posterior-fossa tumor segmentation in MRI. IEEE Transactions on Information Technonogy in Biomedicine. 2011;15(2):206-13.

169. Hongmin C, Verma R, Yangming O, Seung-koo L, Melhem ER, Davatzikos C. Probabilistic segmentation of brain tumors based on multimodality magnetic resonance images. 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro 12-15 April 2007: IEEE; 2007. p. 600-3.

170. Weibei D, Aoyan D, Ping C, Shaowu L, Constans J. Brain tumor segmentation through data fusion of T2-weighted image and MR spectroscopy. 5th International Conference on Bioinformatics and Biomedical Engineering; 10-12 May 2011: IEEE; 2011. p. 1-4.

171. Tinsley HEA, Weiss DJ. Interrater reliability and agreement. In: Howard E.A. Tinsley, Steven D. Brown, editors. Handbook of Applied Multivariate Statistics and Mathematical Modeling. 1st ed: Academic Press; 2000. p. 95-124.

172. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin. 1979;86(2):420-8.

173. Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1993;15(9):850-63.

174. Dubuisson MP, Jain AK. A modified Hausdorff distance for object matching. 12th IAPR International Conference on Computer Vision & Image Processing; 9-13 Oct. 1994; Jerusalem IEEE; 1994. p. 566-8.

175. Jezernik M, Šentjurc M, Schara M. Proton magnetic resonance characterization of brain tumours. Acta Neurochirurgica. 1983;67(1-2):1-9.

176. Hollis DP, Economou JS, Parks LC, Eggleston JC, Saryan LA, Czeisler JL. Nuclear magnetic resonance studies of several experimental and human malignant tumors. Cancer Research. 1973;33(9):2156-60.

177. Kjaer L, Thomsen C, Gjerris F, Mosdal B, Henriksen O. Tissue characterization of intracranial tumors by MR imaging. In vivo evaluation of T1- and T2-relaxation behavior at 1.5 T. Acta Radiologica. 1991;32(6):498-504.

178. Peet AC, Arvanitis TN, Leach MO, Waldman AD. Functional imaging in adult and paediatric brain tumours. Nature Reviews Clinical Oncology. 2012;9:700-11.

179. Forbes JA, Chambless LB, Smith JG, Wushensky CA, Lebow RL, Alvarez J, et al. Use of T2 signal intensity of cerebellar neoplasms in pediatric patients to guide preoperative staging of the neuraxis. Journal of Neurosurgery: Pediatrics. 2011;7(2):165-74.

180. Mahmoud-Ghoneim D, Toussaint G, Constans J, de Certaines J. Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas. Magnetic Resonance Imaging. 2003;21:983 - 7.

181. Langer DL, van der Kwast TH, Evans AJ, Trachtenberg J, Wilson BC, Haider MA. Prostate cancer detection with multi-parametric MRI: logistic regression analysis of quantitative T2,

diffusion-weighted imaging, and dynamic contrast-enhanced MRI. Journal of Magnetic Resonance Imaging. 2009;30(2):327-34.

182. Khayal IS, McKnight TR, McGue C, Vandenberg S, Lamborn KR, Chang SM, et al. Apparent diffusion coefficient and fractional anisotropy of newly diagnosed grade II gliomas. NMR in Biomedicine. 2009;22(4):449-55.

183. Oksuzler YF, Cakmakci H, Kurul S, Oksuzler M, Dirik E. Diagnostic value of diffusion-weighted magnetic resonance imaging in pediatric cerebral diseases. Pediatric Neurology. 2005;32(5):325-33.

184. Fetit AE, Novak J, Rodriguez D, Auer DP, Clark CA, Grundy RG, et al. MRI texture analysis in paediatric oncology: a preliminary study. Studies in Health Technology and Informatics. 2013;190:169-71.

185. Le Bihan D. Looking into the functional architecture of the brain with diffusion MRI. Nature Reviews Neuroscience. 2003;4(6):469-80.

186. Chen HJ, Panigrahy A, Dhall G, Finlay JL, Nelson MD, Jr., Bluml S. Apparent diffusion and fractional anisotropy of diffuse intrinsic brain stem gliomas. AJNR American Journal of Neuroradiology. 2010;31(10):1879-85.

187. Calvar JA, Meli FJ, Romero C, Yánez MLCP, Martinez AR, Lambre H, et al. Characterization of brain tumors by MRS, DWI and Ki-67 labeling index. Journal of Neuro-Oncology. 2005;72(3):273-80.

188. Pavlisa G, Rados M, Pavlisa G, Pavic L, Potocki K, Mayer D. The differences of water diffusion between brain tissue infiltrated by tumor and peritumoral vasogenic edema. Clinical Imaging. 2008;33(2):96-101.

189. Kono K, Inoue Y, Nakayama K, Shakudo M, Morino M, Ohata K, et al. The role of diffusion-weighted imaging in patients with brain tumors. AJNR American Journal of Neuroradiology. 2001;22(6):1081-8.

190. Bulakbasi N, Guvenc I, Onguru O, Erdogan E, Tayfun C, Ucoz T. The added value of the apparent diffusion coefficient calculation to magnetic resonance imaging in the differentiation and grading of malignant brain tumors. Journal of Computer Assisted Tomography. 2004;28(6):735-46.

191. Yamasaki F, Kurisu K, Satoh K, Arita K, Sugiyama K, Ohtaki M, et al. Apparent diffusion coefficient of human brain tumors at MR imaging. Radiology. 2005;235(3):985-91.

192. Rumboldt Z, Camacho DL, Lake D, Welsh CT, Castillo M. Apparent diffusion coefficients for differentiation of cerebellar tumors in children. AJNR American Journal of Neuroradiology. 2006;27(6):1362-9.

193. Tozer DJ, Jager HR, Danchaivijitr N, Benton CE, Tofts PS, Rees JH, et al. Apparent diffusion coefficient histograms may predict low-grade glioma subtype. NMR Biomed. 2007;20(1):49-57.

194. Jaremko JL, Jans LB, Coleman LT, Ditchfield MR. Value and limitations of diffusion-weighted imaging in grading and diagnosis of pediatric posterior fossa tumors. AJNR American Journal of Neuroradiology. 2010;31(9):1613-6.

195. Bull J, Saunders D, Clark C. Discrimination of paediatric brain tumours using apparent diffusion coefficient histograms. European Radiology. 2012;22(2):447-57.

196. Bulakbasi N, Kocaoglu M, Ors F, Tayfun C, Ucoz T. Combination of single-voxel proton MR spectroscopy and apparent diffusion coefficient calculation in the evaluation of common brain tumors. AJNR American Journal of Neuroradiology. 2003;24(2):225-33.

197. Tzika A, Astrakas L, Zarifi M, Petridou N, Young-Poussaint T, Goumnerova L, et al. Multiparametric MR assessment of pediatric brain tumors. Neuroradiology. 2003;45(1):1-10.

198. Costanzo A, Scarabino T, Trojsi F, Giannatempo G, Popolizio T, Catapano D, et al. Multiparametric 3T MR approach to the assessment of cerebral gliomas: tumor extent and malignancy. Neuroradiology. 2006;48(9):622-31.

199. Schneider JF, Confort-Gouny S, Viola A, Le Fur Y, Viout P, Bennathan M, et al. Multiparametric differentiation of posterior fossa tumors in children using diffusion-weighted imaging and short echo-time 1H-MR spectroscopy. Journal of Magnetic Resonance Imaging. 2007;26(6):1390-8.

200. Verma R, Zacharaki EI, Ou Y, Cai H, Chawla S, Lee SK, et al. Multiparametric tissue characterization of brain neoplasms and their recurrence using pattern classification of MR images. Academic Radiology. 2008;15(8):966-77.

201. Bankman IN, Spisz TS, Pavlopoulos S. Two-dimensional shape and texture quantification. Handbook of medical image processing and analysis: Elsevier; 2008. p. 261-77.

202. Haralick RM, Shanmugam K, Dinstein IH. Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics. 1973;3(6):610-21.

203. Soh LK, Tsatsoulis C. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Transactions on Geoscience and Remote Sensing. 1999;37(2):780-95.

204. Weszka JS, Rosenfeld A. An application of texture analysis to materials inspection. Pattern Recognition. 1976;8(4):195-200.

205. Haralick RM. Statistical and structural approaches to texture. Proceedings of the IEEE1979. p. 786-804.

206. Materka A, Strzelecki M. Texture analysis methods - A review. Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels, 1988.

207. Nixon M, Aquado AS. Feature Extraction and Image Processing: Elsevier Newnes; 2002. 1-368 p.

208. Materka A. Texture analysis methodologies for magnetic resonance imaging. Dialogues in Clinical Neuroscience. 2004;6(2):243-50.

209. Cross GR, Jain AK. Markov random field texture models. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1983;5(1):25-39.

210. Geman S, Geman D. Stochastic relaxation, gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1984;PAMI-6(6):721-41.

211. Chaudhuri BB, Sarkar N. Texture segmentation using fractal dimension. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1995;17(1):72-7.

212. Srinivasan GN, Shobha G, editors. Statistical texture analysis. Proceeding of World Academy of Science, Engineering and Technology 2008.

213. Dange A, Khade M, Kulkarni P, Maknikar P. Performance evaluation of different techniques for texture classification 2012.

214. Song-Shan X, Yong-Xing W. Rotation-invariant texture analysis using Radon and Fourier transforms. Journal of Physics. 2007;48(1):1459-64.

215. Gang Z, Ma Z-m. Texture feature extraction and description using gabor wavelet in content-based medical image retrieval. International Conference on Wavelet Analysis and Pattern Recognition; 2-4 Nov. 2007; Beijing: IEEE; 2007. p. 169-73.

216. Popovic M. Texture analysis using 2D wavelet transform: theory and applications. 4th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services; 13-15 Oct 1999; Nis: IEEE; 1999. p. 149-58.

217. De-Shuang H, Xiao-Ping Z, Guang-Bin H, editors. Advances in intelligent computing. International Conference on Intelligent Computing; 2005; Hefei, China: Springer.

218. Lu CS, Chung PC, Chen CF. Unsupervised texture segmentation via wavelet transform. Pattern Recognition. 1997;30(5):729-42.

219. Teuner A, Pichler O, Hosticka BJ. Unsupervised texture segmentation of images using tuned matched Gabor filters. IEEE Transactions on Image Processing. 1995;4(6):863-70.

220. Gabor D. Theory of communication. Part 1: The analysis of information. Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering. 1946;93(26):429-41.

221. Gyaw TA, Ray SR. The wavelet transform as a tool for recognition of biosignals. Biomedical Sciences Instrumentation. 1994;30:63-8.

222. Szczypinski PM, Strzelecki M, Materka A, Klepaczko A. MaZda--A software package for image texture analysis. Computer Methods and Programs in Biomedicine. 2009;94(1):66-76.

223. Weszka JS, Dyer CR, Rosenfeld A. A comparative study of texture measures for terrain classification. IEEE Transactions on Systems, Man and Cybernetics. 1976;SMC-6(4):269-85.

224. Castellano G, Bonilha L, Li LM, Cendes F. Texture analysis of medical images. Clinical Radiology. 2004;59(12):1061-9.

225. Aggarwal N, Agrawal RK. First and second order statistics features for classification of magnetic resonance brain images. Journal of Signal & Information Processing. 2012;3(2):146.

226. Ying L, Viswanath S, Baccon J, Ellison D, Judkins AR, Madabhushi A. A texture-based classifier to discriminate anaplastic from non-anaplastic medulloblastoma. IEEE 37th Annual Northeast Bioengineering Conference; 1-3 April 2011: IEEE; 2011. p. 1-2.

227. John P. Brain tumor classification using wavelet and texture based neural network. International Journal of Scientific & Engineering Research. 2012;3(10).

228. Davies ER. Computer and machine vision: Theory, algorithms, practicalities (Fourth Edition). 4 ed: Elsevier; 2012. p. 209-25.

229. Galloway MM. Texture analysis using gray level run lengths. Computer Graphics and Image Processing. 1975;4(2):172-9.

230. Conners RW, Harlow CA. A theoretical comparison of texture algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1980;PAMI-2(3):204-22.

231. Xiaoou T. Texture information in run-length matrices. IEEE Transactions on Image Processing. 1998;7(11):1602-9.

232. Chu A, Sehgal CM, Greenleaf JF. Use of gray value distribution of run lengths for texture analysis. Pattern Recognition Letters. 1990;11(6):415-9.

233. Dasarathy BV, Holder EB. Image characterizations based on joint gray level-run length distributions. Pattern Recognition Letters. 1991;12(8):497-502.

234. Haar A. Zur Theorie der orthogonalen Funktionensysteme. Mathematische Annalen. 1910;69(3):331-71.

235. Grossman A, Morlet J. Decomposition of Hardy functions into square integrable wavelets of constant shape. SIAM Journal on Mathematical Analysis 1984;15:723-36.

236. Mallat S. Multiresolution approximation and wavelets. Transactions of the American Mathematical Society. 1989;315:69-88.

237. Daubechies I. The wavelet transform, time-frequency localization and signal analysis. IEEE Transactions on Information Theory. 1990;36(5):961-1005.

238. Meyer Y. Ondelettes et opérateurs. Hermann, Paris. 1990;I-III.

239. Mallat SG. Multifrequency channel decompositions of images and wavelet models. IEEE Transactions on Acoustics, Speech and Signal Processing. 1989;37(12):2091-110.

240. Walker JS. A Primer on Wavelets and their scientific applications: Taylor & Francis Group, LLC; 2008.

241. Daubechies I. Ten lectures on wavelet: Rutgers University, New Jersey; 1992.

242. Mohlenkamp MJ, Pereyra MC. Wavelets, their friends, and what they can do for you: European Mathematical Society; 2000.

243. Walker JS. Haar wavelets. 2008. In: A primer on wavelets and their scientific applications [Internet]. Taylor & Francis Group, LLC; [5-40].

244. Zhang X, Chen W. Wavelet domain diffusion for DWI images. 2nd International Conference on Bioinformatics and Biomedical Engineering; 16-18 May 2008; Shanghai IEEE; 2008. p. 2149-52.

245. Elfouly FH, Mahmoud MI, Dessouky MIM, Deyab S. Comparison between Haar and Daubechies wavelet transformations on FPGA technology. International Journal of Computer, Information & Systems Science & Engineering. 2008;2(1):37-42.

246. Wahid K. Low complexity implementation of Daubechies wavelets for medical imaging applications. Discrete Wavelet Transforms - Algorithms and Applications: InTech; 2011.

247. Gao RX, Yan R. Discrete Wavelet Transform. Wavelets: Theory and Applications for Manufacturing: Springer; 2011. p. 63.

248. Kumari S, Vijay R. Effect of symlet filter order on denoising of still images. Advanced Computing: An International Journal. 2012;3(1):137-43.

249. Meyer-Bäse A. Chapter II Feature selection and extraction. Pattern recognition for medical imaging Amsterdam ; Boston : Elsevier Academic Press; 2004. p. 18.

250. Donoho DL, editor High-dimensional data analysis: the curses and blessings of dimensionality. American Mathematical Society Conference Math Challenges of the 21st Century; 2000.

251. Bellman R. Adaptive Control Processes: A Guided Tour: Princeton University Press; 1961.

252. Peng Y, Wu Z, Jiang J. A novel feature selection approach for biomedical data classification. Journal of Biomedical Informatics. 2010;43(1):15-23.

253. Smialowski P, Frishman D, Kramer S. Pitfalls of supervised feature selection. Bioinformatics. 2010;26(3):440-3.

254. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. Journal of Bioinformatics and Computational Biology. 2005;3(2):185-205.

255. Lijun Z, Chun C, Jiajun B, Xiaofei H. A unified feature and instance selection framework using optimum experimental design. IEEE Transactions on Image Processing. 2012;21(5):2379-88.

256. Guyon I, Elisseeff A. An introduction to variable and feature selection. Journal of Machine Learning Research. 2003;3:1157-82.

257. Peng H, Fulmi L, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2005;27(8):1226-38.

258. Zhang Y, Dong Z, Wu L, Wang S. A hybrid method for MRI brain image classification. Expert Systems with Applications. 2011;38(8):10049-53.

259. Polat D, Cataltepe Z. Feature selection and classification on brain computer interface (BCI) data. 20th Signal Processing and Communications Applications Conference; 18-20 April 2012; Mugla: IEEE; 2012. p. 1-4.

260. Tourassi GD, Frederick ED, Markey MK, Floyd CE, Jr. Application of the mutual information criterion for feature selection in computer-aided diagnosis. Medical Physics. 2001;28(12):2394-402.

261. Krier C, François D, Wertz V, Verleysen M. Feature scoring by mutual information for classification of mass spectra. In: D. Ruan, P. D'Hondt, P.F. Fantoni, M. De Cock, M. Nechtegael, Kerre EE, editors. 7th International FLINS Conference on Applied Artificial Intelligence; 29-31 August 2006; Genova, Italy: World Scientific Publishing; 2006. p. 557-64.

262. Andronesi OC, Blekas KD, Mintzopoulos D, Astrakas L, Black PM, Tzika AA. Molecular classification of brain tumor biopsies using solid-state magic angle spinning proton magnetic resonance spectroscopy and robust classifiers. International Journal of Oncology. 2008;33(5):1017-25.

263. Jafari M, Shafaghi R. A hybrid approach for automatic tumor detection of brain MRI using support vector machine and genetic algorithm. Global Journal of Science, Engineering and Technology. 2012(3):1-8.

264. Zhou Y, Yu F, Duong T. Multiparametric MRI characterization and prediction in autism spectrum disorder using graph theory and machine learning. PLoS ONE. 2014;9(6).

265. Gulgezen G, Cataltepe Z, Yu L. Stable and accurate feature selection. In: Buntine W, Grobelnik M, Mladenić D, Shawe-Taylor J, editors. Machine Learning and Knowledge Discovery in Databases. Lecture Notes in Computer Science. 5781: Springer Berlin Heidelberg; 2009. p. 455-68.

266. Iftekharuddin KM, Zheng J, Islam MA, Ogg RJ, Lanningham F. Brain tumor detection in MRI: Technique and statistical validation. Fortieth Asilomar Conference on Signals, Systems and Computers; 29 Oct. 2006-1 Nov. 2006; Pacific Grove, CA: IEEE; 2006. p. 1983-7.

267. Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, Melhem ER, et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magnetic Resonance in Medicine. 2009;62(6):1609-18.

268. Othman MF, Basri MA. Probabilistic neural network for brain tumor classification. 2nd International Conference on Intelligent Systems, Modelling and Simulation; 25-27 Jan. 2011: IEEE; 2011. p. 136-8.

269. Cortes C, Vapnik V. Support-Vector Networks. Machine Learning. 1995;20(3):273-97.

270. Cristianini N, Shawe-Taylor J. An introduction to Support Vector Machines: and other kernel-based learning methods: Cambridge : Cambridge University Press 2000.

271. Chang C-C, Lin C-J. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology. 2011;2(3):1-27.

272. Nelder JA, Wedderburn RWM. Generalized Linear Models. Journal of the Royal Statistical Society Series A (General). 1972;135(3):370-84.

273. Burns RB, Burns RA. Business research methods and statistics using SPSS: SAGE Publications; 2008.

274. Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression: John Wiley & Sons; 2013.

275. Chen J, Shao J. Iterative Weighted Least Squares Estimators. The Annals of Statistics. 1993;21(2):1071-92.

276. Peng C-YJ, Lee KL, Ingersoll GM. An introduction to logistic regression analysis and reporting. The Journal of Educational Research. 2002;96(1):3-14.

277. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. Bioinformatics. 2005;21(15):3301-7.

278. Ambroise C, McLachlan GJ, editors. Selection bias in gene extraction on the basis of microarray gene-expression data. Proceedings of the National Academy of Sciences of the United States of America; 2002 14 May 2002.

279. Teh SY, Othman AR, Khoo MBC, editors. Dichotomous logistic regression with leave-one-out validation. World Academy of Science, Engineering and Technology 38; 2010.

280. Azuaje F. Genomic data sampling and its effect on classification performance assessment. BMC Bioinformatics. 2003;4(1):5.

281. Breiman L, Spector P. Submodel Selection and Evaluation in Regression. The X-Random Case. International Statistical Review / Revue Internationale de Statistique. 1992;60(3):291-319.

282. Efron B, Tibshirani R. Improvements on Cross-Validation: The .632+ Bootstrap Method. Journal of the American Statistical Association. 1997;92(438):548-60.

283. Efron B. Estimating the error rate of a prediction rule: Improvement on cross-validation. Journal of the American Statistical Association. 1983;78(382):316-31.

284. Bishop CM. Pattern recognition and machine learning: New York, N.Y. : Springer 2006.

285. Davies NP, Wilson M, Harris LM, Natarajan K, Lateef S, Macpherson L, et al. Identification and characterisation of childhood cerebellar tumours by in vivo proton MRS. NMR in Biomedicine. 2008;21(8):908-18.

286. Fleiss JL, Levin B, Paik MC. Statistical methods for rates and proportions. 3rd ed. Balding DJ, Cressie NAC, Fisher NI, Johnstone IM, Kadane JB, Ryan LM, et al., editors: John Wiley & Sons, Inc.; 2003.

287. Foody GM. Classification accuracy comparison: Hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. Remote Sensing of Environment. 2009;113(8):1658-63.

288. Behrens TEJ, Woolrich MW, Jenkinson M, Johansen-Berg H, Nunes RG, Clare S, et al. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magnetic Resonance in Medicine. 2003;50(5):1077-88.

289. Zhang Y, Zhu H, Mitchell JR, Costello F, Metz LM. T2 MRI texture analysis is a sensitive measure of tissue injury and recovery resulting from acute inflammatory lesions in multiple sclerosis. NeuroImage. 2009;47(1):107-11.

290. He L, Peng Z, Everding B, Wang X, Han CY, Weiss KL, et al. A comparative study of deformable contour methods on medical image segmentation. Image and Vision Computing. 2008;26(2):141-63.

291. Likert R. A technique for the measurement of attitudes. Archives of Psychology. 1932;22(140):1-55.

292. Ekin A. Pathology-robustmr intensity normalization with global and local constraints. IEEE International Symposium on Biomedical Imaging: From Nano to Macro; 30 March 2011-2 April 2011; Chicago, IL: IEEE; 2011. p. 333-6.

293. Bauer S, Wiest R, Nolte LP, Reyes M. A survey of MRI-based medical image analysis for brain tumor studies. Physics in Medicine and Biology. 2013;58(13):R97-129.

294. Aprile I, Iaiza F, Lavaroni A, Budai R, Dolso P, Scott CA, et al. Analysis of cystic intracranial lesions performed with fluid-attenuated inversion recovery MR imaging. AJNR American Journal of Neuroradiology. 1999;20(7):1259-67.

295. Khademi A, Krishnan S, Venetsanopoulos A. Shift-Invariant DWT for Medical Image Classification In: Olkkonen JT, editor. Discrete Wavelet Transforms-Theory and Applications: InTech; 2011.

296. Olkkonen H, Olkkonen JT. Shift invariant discrete wavelet transforms. Discrete Wavelet Transforms - Algorithms and Applications: InTech; 2011.

297. Zhang J, Tan T. Brief review of invariant texture analysis methods. Pattern Recognition. 2002;35(3):735-47.

298. Pantelis G. Computer assisted diagnosis of brain tumors based on statistical methods and pattern recognition techniques: University of Patras; 2010.

299. Sinha S, Bastin ME, Whittle IR, Wardlaw JM. Diffusion tensor MR imaging of high-grade cerebral gliomas. AJNR American Journal of Neuroradiology. 2002;23(4):520-7.

300. Pope W, Qiao X, Kim H, Lai A, Nghiemphu P, Xue X, et al. Apparent diffusion coefficient histogram analysis stratifies progression-free and overall survival in patients with recurrent GBM treated with bevacizumab: a multi-center study. Journal of Neuro-Oncology. 2012;108(3):491-8.

301. Miao H, Fukatsu H, Ishigaki T. Prostate cancer detection with 3-T MRI: Comparison of diffusion-weighted and T2-weighted imaging. European Journal of Radiology. 2007;61(2):297-302.

302. Savio SJ, Harrison LC, Luukkaala T, Heinonen T, Dastidar P, Soimakallio S, et al. Effect of slice thickness on brain magnetic resonance image texture analysis. Biomedical Engineering Online. 2010;9:60.

303. Clausi DA, Zhao Y. Grey level co-occurrence integrated algorithm (GLCIA): a superior computational method to rapidly determine co-occurrence probability texture features. Computers & Geosciences. 2003;29(7):837-50.

304. Yang X, Tridandapani S, Beitler JJ, Yu DS, Yoshida EJ, Curran WJ, et al. Ultrasound GLCM texture analysis of radiation-induced parotid-gland injury in head-and-neck cancer radiotherapy: an in vivo study of late toxicity. Medical Physics. 2012;39(9):5732-9.

305. Baraldi A, Parmiggiani F. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. IEEE Transactions on Biomedical Engineering. 1995;33(2):293-304.

306. Linfoot EH. An informational measure of correlation. Information and Control. 1957;1(1):85-9.