

**AN AUTOMATED SYSTEM FOR THE
CLASSIFICATION AND SEGMENTATION OF BRAIN
TUMOURS IN MRI IMAGES BASED ON THE
MODIFIED GREY LEVEL CO-OCCURRENCE
MATRIX**

Ali Majeed Hasan Al-Waeli

Thesis Submitted in Partial Fulfilment of the Requirements of
the Degree of Doctor of Philosophy

School of Computing, Science and Engineering
University of Salford, Salford, UK

2017

TABLE OF CONTENTS

TABLE OF CONTENTS	I
LIST OF TABLES	V
LIST OF FIGURES	VI
ACKNOWLEDGMENT	XI
DECLARATION	XII
ABBREVIATIONS	XIII
LIST OF PUBLICATIONS	XVI
ABSTRACT	XVII
CHAPTER ONE: Introduction	1
1.1 Introduction.....	1
1.2 Research Questions.....	6
1.3 Aim and Objectives.....	6
1.4 Contribution of the Study.....	7
1.5 Research Methodology	8
1.6 Thesis Organization	10
CHAPTER TWO: Literature Review	13
2.1 Introduction.....	13
2.2 Pre-processing Analysis.....	15
2.2.1 Image Enhancement.....	15
2.2.2 MRI Intensity Normalization	16
2.2.3 Mid-Sagittal Plane of Brain Detection.....	17
2.3 Texture Analysis	18
2.3.1 Grey Level Co-occurrence Matrix	22
2.3.2 Gabor Wavelet.....	24
2.4 Preparing the Extracted Features for Classification.....	25
2.4.1 Relevance Analysis	25
2.4.1.1 One-Way Analysis of Variance.....	27
2.4.1.2 Principle Component Analysis	28
2.4.2 Features Transformation	29
2.4.2.1 Min-Max Normalization.....	30
2.4.2.2 Z-Score Normalization	30
2.4.2.3 Decimal Scaling Normalization.....	31

2.4.2.4 Median Normalization.....	31
2.4.2.5 Sigmoid Normalization.....	31
2.5 Classification.....	32
2.5.1 Linear Discriminant Analysis.....	32
2.5.2 Support Vector Machines.....	34
2.5.3 Artificial Neural Network	37
2.5.3.1 Multilayer Perceptron	38
2.5.3.2 Determining Number of Neurons in the Hidden Layers of MLP	40
2.6 Image Segmentation.....	41
2.6.1 Pixel Based Segmentation.....	44
2.6.2 Region Based Segmentation.....	45
2.6.3 Edge Based Segmentation.....	46
2.6.4 Deformable Model	46
2.6.4.1 Active Contour	48
2.6.4.2 Level Set Method.....	51
2.6.4.3 2D Active Contour without Edge	54
2.6.4.4 3D Active Contour without Edge	58
2.7 Conclusion	60
CHAPTER THREE: Basics of Magnetic Resonance Imaging	61
3.1 Introduction.....	61
3.2 Digital Image Representation	61
3.3 Magnetic Resonance Imaging.....	62
3.4 Al-Kadhimiya Teaching Hospital.....	70
3.5 Brain Tumours	71
3.6 Conclusion	73
CHAPTER FOUR: Image Pre-processing Analysis of MRI	74
4.1 Introduction.....	74
4.2 Data Collection	74
4.3 Resizing the Dimensions of MRI slices.....	77
4.4 MRI Enhancement Algorithm.....	78
4.5 Intensity Normalization	79
4.6 Background Segmentation.....	80
4.7 MSP Detection and Correction	82
4.7.1 Orientation Determination Based PCA	83

4.7.2 Geometrical Transformation of Patient’s Head	86
4.7.3 Centralize Patient’s Head in the Centre of MRI Slice	88
4.8 Conclusion	92
CHAPTER FIVE: Features Extraction and Classification	93
5.1 Introduction.....	93
5.2 Preparing MRI Brain Slices for Feature Extraction.....	94
5.2.1 Image Cropping.....	94
5.2.2 Image Resizing.....	95
5.2.3 Modified Grey Level Co-occurrence Matrix Method	95
5.3 Texture Descriptions	98
5.3.1 Contrast	99
5.3.2 Correlation.....	99
5.3.3 Entropy	100
5.3.4 Energy	100
5.3.5 Homogeneity	101
5.3.6 Dissimilarity	101
5.3.7 Sum of Square Variance.....	101
5.3.8 Cluster Shade.....	102
5.3.9 Cluster Prominence	102
5.3.10 Inverse Difference Normalized	102
5.3.11 Inverse Difference Moment Normalized.....	103
5.3.12 Sum Average, Sum Entropy, Sum Variance and Difference Entropy	103
5.3.13 Information Measure of Correlation I and Correlation II	104
5.3.14 Autocorrelation.....	105
5.3.15 Maximum Probability	105
5.3.16 Weighted Mean	105
5.3.17 Weighted Distance	106
5.3.18 Cross-Correlation Coefficients.....	107
5.4 Feature Aggregation.....	108
5.5 Feature Selection.....	108
5.6 Feature Normalization	109
5.7 Feature Classification.....	109
5.8 Conclusion	122
CHAPTER SIX: Three-Dimensional Modified Grey Level Co-occurrence Matrix	124

6.1 Introduction.....	124
6.2 Three-Dimensional Modified Grey Level Co-occurrence Matrix	125
6.3 Feature Aggregation.....	128
6.4 Feature Selection.....	129
6.5 Feature Normalization	129
6.6 Feature Classification.....	129
6.7 Conclusion	135
CHAPTER SEVEN: Detection of the Tumour Location and Slices	136
7.1 Introduction.....	136
7.2 Overview of Genetic Algorithms.....	136
7.3 Termination of GA.....	143
7.4 Bounding 3D-Boxes Based Genetic Algorithm Method	143
7.4.1 Exponential Transformation of MRI Brain Slices	144
7.4.2 Designing of GA	145
7.4.2.1 Individual Construction.....	145
7.4.2.2 Population Size	147
7.5 Conclusion	151
CHAPTER EIGHT: Brain Tumour Segmentation	153
8.1 Introduction.....	153
8.2 Brain Tumour Segmentation.....	156
8.2.1 2DACWE Segmentation Outcome	156
8.2.2 3DACWE Segmentation Outcome	157
8.3 Conclusion	168
CHAPTER NINE: Conclusion and Future Work	171
9.1 Introduction.....	171
9.2 Future Works	173
REFERENCES.....	175
APPENDIX A	188

LIST OF TABLES

Table 2.1: Summary of existing MSP methods.....	18
Table 2.2: List of the 19 co-occurrence statistics extracted from GLCM.....	23
Table 4.1: The results of predicted yaw angle.....	91
Table 5.1: Comparison of MRI brain scans features (mean \pm standard deviation (SD)) between normal and abnormal patients.....	110
Table 5.2: The results of 10 runs for 190 input descriptors with different number of neurons in hidden layer.....	113
Table 5.3: The result of using four different F-statistic threshold values on the classification accuracy.....	115
Table 5.4: The results of 10 runs for 100 input descriptors with different number of neurons in hidden layer.....	116
Table 5.5: Classification accuracy for the three classifiers with PCA method for feature selection.....	120
Table 5.6: Classification accuracy for the three classifiers with Kernel PCA method for feature selection.....	120
Table 5.7: Classification accuracy for the three classifiers with WEKA CFS-SE method for feature selection.....	120
Table 5.8: Comparison with previous proposed methods.....	121
Table 6.1: Comparison of MRI brain scans feature (mean \pm standard deviation (SD)) between normal and abnormal patients.....	130
Table 6.2: The result of using four different F-statistic threshold values on the classification accuracy.....	132
Table 6.3: The results of 10 runs for 100 input descriptors with different number of neurons in hidden layer.....	133
Table 7.1: An average number of iterations with different population size and mutation rates.....	148
Table 7.2: A comparison between clinical and experimental slice identification.....	151
Table 8.1: Segmentation results for each patient in the collected dataset.....	164
Table 8.2: Segmentation results for each patient in the BRATS 2013 dataset.....	165
Table 8.3: A comparison between clinical and experimental slice identification.....	166
Table 8.4: An overview of segmentation methods compared to proposed system.....	169

LIST OF FIGURES

Figure 1.1: The major polluted sites with DU in Iraq (Al Hilfi et al., 2013).....	2
Figure 1.2: Samples of four pathological MRI slices: A) T2-w, B) T1-w, C) FLAIR and D) T1c-w.....	4
Figure 1.3: The research phases of the proposed system.....	11
Figure 2.1: A general framework of the screening system.....	14
Figure 2.2: Representation of GLCM.....	23
Figure 2.3: Implementation of GLCM; A) Original Matrix, B) The four GLCM with $d=1$ and orientations ($\theta =0^\circ, 45^\circ, 90^\circ$ and 135°).....	24
Figure 2.4: Separation between two classes.....	34
Figure 2.5: The discriminant distributions.....	34
Figure 2.6: Two classes separated by a hyperplane.....	35
Figure 2.7: Multilayer perceptron configuration (Hu and Hwang, 2001).....	39
Figure 2.8: Evolving of contour C	51
Figure 3.1: The representation of digital image.....	62
Figure 3.2: MRI brain slices with 5.5 mm thickness.....	63
Figure 3.3: Internal rotation of a hydrogen proton.....	64
Figure 3.4: Random oriented poles of hydrogen protons.....	64
Figure 3.5: Aligned poles of hydrogen protons in external magnetic field.....	65
Figure 3.6: Direction of the net magnetization.....	66
Figure 3.7: The precession movement of the proton.....	66
Figure 3.8: The net magnetization tilts toward the x-y plane.....	67
Figure 3.9: The net magnetization is relaxed after switching off the RF pulse.....	67
Figure 3.10: Excitation and relaxation of the hydrogen protons (Petrou, 2011), A) without excitation, B) with excitation, C,D,E) T2 relaxation (Spin-Spin Relaxation), F, G) T1 relaxation (Spin-Lattice Relaxation) and H) relaxation complete.....	69
Figure 3.11: Al-Kadhimiya Teaching Hospital.....	71
Figure 4.1: Overall flow chart of the proposed system.....	75
Figure 4.2: SIEMENS MAGNETOM Avanto MRI scanner in MRI Unit of Al Kadhimiya Teaching Hospital- Iraq.....	76
Figure 4.3: Examples from the used datasets: The collected dataset in the first row and the standard dataset in the second row.....	77
Figure 4.4: MRI slices resizing process, A) Original MRI slice with dimensions 512×470 , B) Resized MRI slice with dimensions 512×512	78

Figure 4.5: MRI brain scan image, A) Original image, B) Filtered image using Gaussian filter with ($\sigma=0.5$), C) Filtered image using Gaussian filter with ($\sigma=1$), D) Filtered image using Gaussian filter with ($\sigma=2$).....	79
Figure 4.6: Histogram normalization method implementation, A) Original MRI slice, B) Normalized MRI slice.....	80
Figure 4.7: T2-w images of two different patients and corresponding histograms.....	81
Figure 4.8: An example of skull boundary identification, A) Original MRI slice, B) Segmented MRI slice with threshold equal to 25, C) Dilated MRI slice and D) Filled holes in MRI slice.....	82
Figure 4.9: How the patient's head lies in the MRI head coil (UC Davis Medical Center, 2016).....	84
Figure 4.10: Remapping the axes (X,Y) of the original data into new axes (X',Y').....	84
Figure 4.11: Original and new coordinates of brain.....	86
Figure 4.12: Mapping pixels by nearest neighbour interpolation.....	87
Figure 4.13: Mapping pixels by bilinear interpolation.....	88
Figure 4.14: Steps for detecting and correcting the MSP of the brain in MR brain slices....	90
Figure 4.15: The solid yellow lines denote the computed MSPs and the dashed red lines denote the clinicians delineation.....	90
Figure 4.16: Distribution of MSE between manual and our algorithm delineation of MSPs.....	91
Figure 4.17: Resampling of one slice from the axial MRI brain scanning image with varied rotate angles.....	91
Figure 5.1: Flowchart of implementing feature extraction, selection and classification....	94
Figure 5.2: MRI brain slice cropping.....	95
Figure 5.3: How reference pixel relates with opposite nine pixels.....	96
Figure 5.4: A) MRI normal brain slice, B) MRI abnormal brain slice, C) MGLCM of normal brain scanning and D) MGLCM of abnormal brain scanning.....	98
Figure 5.5: Weighted mean representation on MGLCM.....	106
Figure 5.6: Weighted distance representation on MGLCM.....	107
Figure 5.7: Flowchart and steps of training and testing of classification techniques.....	112
Figure 5.8: Number of neurons vs. the mean of RMSE of runs.....	114
Figure 5.9: A) MLP network structure and B) The performance of MLP network.....	114
Figure 5.10: A comparison between MGLCM, GLCM and Gabor wavelet regards classification accuracy.....	115
Figure 5.11: The optimal number of features corresponding with classification accuracy..	116

Figure 5.12: Number of neurons vs. RMSEs of runs.....	117
Figure 5.13: A) MLP network structure and B) The performance of MLP network.....	118
Figure 5.14: Textural features (mean \pm standard error (SE)) of the normal and pathological MRI brain scans.....	119
Figure 5.15: A comparison of classification accuracies for MGLCM, GLCM and Gabor wavelet using ANOVA method.....	119
Figure 5.16: Performance of the implemented feature selection methods.....	121
Figure 6.1: Flowchart of the implementation of the three-dimensional feature extraction, selection and classification.....	125
Figure 6.2: How reference pixel relates with opposite nine pixels.....	126
Figure 6.3: A) MRI abnormal brain scanning image, B) MRI normal brain scanning image, C) 3DMGLCM of abnormal brain scanning and D) 3DMGLCM of normal brain scanning.....	128
Figure 6.4: Flowchart and steps of training and testing of classification techniques.....	131
Figure 6.5: A comparison of classification accuracies for 3DMGLCM, MGLCM, Gabor wavelet and GLCM.....	131
Figure 6.6: The optimal number of features corresponding with classification accuracy..	132
Figure 6.7: Number of neurons vs. RMSEs of runs.....	133
Figure 6.8: A) MLP network structure and B) The performance of MLP network.....	134
Figure 6.9: Textural features (mean \pm standard error (SE)) of the normal and pathological MRI brain scans.....	135
Figure 7.1: Flowchart of GA.....	138
Figure 7.2: How the variables are organized in the chromosome.....	139
Figure 7.3: Roulette wheel selection method.....	140
Figure 7.4: Stochastic universal selection method.....	140
Figure 7.5: Single point crossover operation.....	141
Figure 7.6: Multi-points crossover operation.....	142
Figure 7.7: Single point mutation operation.....	142
Figure 7.8: Image exponential transformation, A) Original MRI brain image, B) Transformed MRI brain image.....	144
Figure 7.9: Pseudo-code for BBBGA.....	145
Figure 7.10: Representation of one 3D-box in the left hemisphere of brain using $(x_1, x_2, y_1, y_2, z_1, z_2)$ coordinates and opposite region.....	146
Figure 7.11: Individual structure.....	146
Figure 7.12: Average of GA iterations for different values of population size and mutation rate.....	148

Figure 7.13: MRI brain slices: the output of the BBBGA implementation on pathological MRI brain slices with a population size equal to 100 and a mutation rate equal to 0.05. The red rectangles that denotes the optimized 3D-box stands on pathological area in slices 6 to 9, where the tumour are appeared.....	149
Figure 7.14: RMSE decreases within 18 iterations by GA.....	150
Figure 7.15: FP increases with increasing noise amount in the MRI scan.....	151
Figure 8.1: Comparison of a binary segmentation image (Seg.) with the reference image (Ref.) and the statistical measures.....	155
Figure 8.2: Comparative segmentation results on T2-w MRI slices by 2DACWE. The ground truth is marked in green and the output of 3DACWE is marked in red.....	157
Figure 8.3: Comparative segmentation results on T2-w MRI slices by 3DACWE. The ground truth is marked in green and the output of 3DACWE is marked in red.....	158
Figure 8.4: The Chan-Vese energy function convergence to steady state.....	158
Figure 8.5: Comparative segmentation results on T1-w MRI slices by 3DACWE. The ground truth is marked in green and the output of 3DACWE is marked in red.....	159
Figure 8.6: The Chan-Vese energy function convergence to steady state.....	160
Figure 8.7: Comparative segmentation results on FLAIR slices by 3DACWE. The ground truth is provided by BRATS 2013 in green and the output of 3DACWE is marked in red.....	161
Figure 8.8: The Chan-Vese energy function convergence to steady state.....	162
Figure 8.9: Comparative segmentation results on T1-w slices by 3DACWE. The ground truth is marked in green and the output of 3DACWE is marked in red.....	163
Figure 8.10: The Chan-Vese energy function convergence to steady state.....	164
Figure 8.11: The overall results of the segmentation of the four MRI modalities.....	166
Figure 8.12: Comparing tumour segmentation average Dice scores of four modalities of MRI using 2DACWE and 3DACWE.....	167
Figure 8.13: Comparison between 3DACWE and 2DACWE segmentation results for collected dataset (50 patients).....	168
Figure 8.14: Scatter plot of Dice score to the summation of slice thickness and space between slices, showing mean of Dice score (R2) as the dotted red line.....	170
Figure A.1: Result of the BBBGA implementation on MRI brain scan of a pathological patient when there is no additive Gaussian noise ($\sigma=0$).....	188
Figure A.2: Result of the BBBGA implementation on MRI brain scan of a pathological patient when there is additive Gaussian noise ($\sigma=0.1$).....	189

Figure A.3: Result of the BBBGA implementation on MRI brain scan of a pathological patient when there is additive Gaussian noise ($\sigma=0.2$).....	189
Figure A.4: Result of the BBBGA implementation on MRI brain scan of a pathological patient when there is additive Gaussian noise ($\sigma=0.3$).....	190
Figure A.5: Result of the BBBGA implementation on MRI brain scan of a pathological patient when there is additive Gaussian noise ($\sigma=0.4$).....	190

ACKNOWLEDGMENT

First and foremost, I would like to express my sincere gratitude to the ministry of higher education and scientific research, Al-Nahrain University-College of Medicine in Iraq, for giving me the opportunity to complete my PhD study in the universities of United Kingdom.

I would like to express my deepest appreciation to my supervisors Prof. Farid Meziane and Dr. Rob Aspin for their invaluable contributions, encouragement and guidance through the period of my study. Also, I would like to thank Dr. Hamid Abdulla Jalab for providing help whenever needed with publications.

I would like to thank the Magnetic Resonance Imaging Unit of Al-Kadhimiya Teaching Hospital in Iraq and all of its staff members for their support and kindness throughout this work.

Lastly, I wish to thank my parents, my wife, siblings and my dearest daughters Yusur and Fatima for their support.

DECLARATION

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university, or institute of learning.

ABBREVIATIONS

2DACWE	Two-Dimensional Active Contour without Edge
2DDWT	Two-Dimensional Discrete Wavelet Transform
3DACWE	Three-Dimensional Active Contour without Edge
AC	Active Contour
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
BBGGA	Bounding 3D-boxes based genetic algorithm
BPNN	Back Propagation Neural Network
BRATS 2013	Standard Benchmark Multimodal Brain Tumour Segmentation 2013
BUS	Breast Ultrasound Images
CBAC	Content Based Active Contour
CBTRUS	Central Brain Tumour Registry of the United States
CFS-SE	Correlation Feature Selection Subset Evaluator
cm ³	Cubic centimetre
CNN	Convolutional Neural Network
<i>cov</i>	Covariance Matrix
CSF	Cerebrospinal Fluid
CT	Computed Tomography
DCT	Discrete Cosine Transform
DU	Depleted Uranium
FLAIR	Fluid Attenuated Inversion Recovery
FN	False Negative
FP	False Positive
GA	Genetic Algorithm
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
GE	Gradient Echo
GLCM	Grey Level Co-occurrence Matrix
GLRLM	Grey Level Run Length Matrix
GM	Grey Matter

GTR	Gross Total Resection
HOG	Histogram of Oriented Gradient
IBF	Intensity-Based Features
IDMN	Inverse Difference Moment Normalized
IDN	Inverse Difference Normalized
IR	Inversion Recovery
IRe	Image Resolution
Kernel PCA	Kernel Principle Component Analysis
KFCM	Kernel-Based Fuzzy C-Means
KNN	K-Nearest Neighbour
LBP	Local Binary Pattern
LDA	Linear Discriminant Analysis
LoG	Laplacian of Gaussian
MGLCM	Modified Grey Level Co-occurrence Matrix
MI	Mutual Information
MLP	Multi-layer Perceptron Neural Network
mm	millimetre
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
ms	Millisecond
MSE	Mean Squared Error
MSP	Mid-Sagittal Plane
PCA	Principal Component Analysis
PCs	Principal Components
PV	Partial Volumes
RAM	Random Access Memory
Ref.	Reference
RF	Radio Frequency
RILBP	Rotation Invariant Local Binary Patterns
RMSE	Root Mean Square Error
RoI	Region of Interest

sec.	Second
SBF	Shape Based Feature
SD	Standard Deviation
SDF	Signed Distance Function
SE	Spin Echo
Seg.	Segmentation
SVM	Support Vector Machine
T1c-w	T1-weighted images with contrast enhancement
T1-w	T1-weighted image
T2-w	T2-weighted image
TE	Echo Time
TN	True Negative
TP	True Positive
TR	Repetition Time
US	Ultrasonography
WM	White Matter

LIST OF PUBLICATIONS

1. A. Hasan and T. Ritchings. Automated segmentation and visualization of tumours in MRI brain scans, Salford Postgraduate Annual Research Conference (SPARC 2015), University of Salford, UK, Page 100.
2. A. Hasan, F. Meziane and M. Kadhim. Automated Segmentation of Tumours in MRI Brain Scans. In Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC), 2016, Rome, Italy. SCITEPRESS, Pages 55-62.
3. A. Hasan and F. Meziane (2016). Automated screening of MRI brain scanning using grey level statistics, Computers & Electrical Engineering, Volume 53, July 2016, Pages 276–291.
4. A. Hasan, F. Meziane and R. Aspin. Automated Screening and Segmentation of MRI Brain Scanning Based on 3-Dimensional Active Contour without Edge, Salford Postgraduate Annual Research Conference (SPARC 2016), University of Salford, UK, Page 34.
5. A. Hasan, F. Meziane and H. Jalab. Performance of Grey Level Statistic Features versus Gabor wavelet for Screening MRI Brain Tumors: A Comparative Study, International Conference on Information Communication and Management (ICICM 2016), IEEE. University of Hertfordshire, UK. Pages 136-140.
6. A. Hasan, F. Meziane, R. Aspin and Hamid A. Jalab (2016). Segmentation of Brain Tumors in MRI Images Using Three-Dimensional Active Contour without Edge, Symmetry, 8(11), Page 132; doi:10.3390/sym8110132.
7. A. Hasan, F. Meziane, R. Aspin and H. Jalab. MRI Brain Scan Classification Using Novel 3-D Statistical Features, Accepted in 2nd International Conference on Internet of Things, Data and Cloud Computing (ICC 2017), University of Cambridge, UK.

ABSTRACT

The development of an automated system for the classification and segmentation of brain tumours in MRI scans remains challenging due to high variability and complexity of the brain tumours. Visual examination of MRI scans to diagnose brain tumours is the accepted standard. However due to the large number of MRI slices that are produced for each patient this is becoming a time consuming and slow process that is also prone to errors.

This study explores an automated system for the classification and segmentation of brain tumours in MRI scans based on texture feature extraction. The research investigates an appropriate technique for feature extraction and development of a three-dimensional segmentation method. This was achieved by the investigation and integration of several image processing methods that are related to texture features and segmentation of MRI brain scans. First, the MRI brain scans were pre-processed by image enhancement, intensity normalization, background segmentation and correcting the mid-sagittal plane (MSP) of the brain for any possible skewness in the patient's head. Second, the texture features were extracted using modified grey level co-occurrence matrix (MGLCM) from T2-weighted (T2-w) MRI slices and classified into normal and abnormal using multi-layer perceptron neural network (MLP). The texture feature extraction method starts from the standpoint that the human brain structure is approximately symmetric around the MSP of the brain. The extracted features measure the degree of symmetry between the left and right hemispheres of the brain, which are used to detect the abnormalities in the brain. This will enable clinicians to reject the MRI brain scans of the patients who have normal brain quickly and focusing on those who have pathological brain features. Finally, the bounding 3D-boxes based genetic algorithm (BBBGA) was used to identify the location of the brain tumour and segments it automatically by using three-dimensional active contour without edge (3DACWE) method.

The research was validated using two datasets; a real dataset that was collected from the MRI Unit in Al-Kadhimiya Teaching Hospital in Iraq in 2014 and the standard benchmark multimodal brain tumour segmentation (BRATS 2013) dataset.

The experimental results on both datasets proved that the efficacy of the proposed system in the successful classification and segmentation of the brain tumours in MRI scans. The achieved

classification accuracies were 97.8% for the collected dataset and 98.6% for the standard dataset. While the segmentation's Dice scores were 89% for the collected dataset and 89.3% for the standard dataset.

CHAPTER ONE

Introduction

Overview

This chapter provides an introduction to the research conducted in this study followed by the research questions and the aim and objectives. The chapter concludes with asserting the contributions of the study and the research methodology used. Finally, a brief description of the remainder chapters of the thesis is outlined.

1.1 Introduction

Medical imaging is a powerful technology used for gaining an insight into medical abnormalities by creating visual representations of the internal organs or tissues of the body and is used for clinical diagnosis. Image processing has been embedded in medical systems and applications and is used now widely in medicine from diagnosis to therapy. It has a significant impact on the performance of digital medical diagnostic workflows and the clinicians who use them. The auto classification and segmentation of brain tumours has potential to further improve the accuracy of diagnosis of all acquired medical images.

In 2004, the Iraqi Ministry of Health reported that the average annual number of registered cancerous tumour cases between 1995 and 1997 were 8000 to 9000 cases. This rate rose to approximately 11,000 in 2000. The majority of these cases were from the south and the middle of Iraq (Fathi et al., 2013; Alwan, 2004). These regions were the battlefields during the First and Second Gulf Wars in 1990 and 1997 respectively and led to significant environmental pollution. There are more than 350 sites that were polluted with Depleted Uranium (DU) and other toxins as shown in Fig. 1.1. In addition to the impact of sanctions and occupation, Iraq medical health services have completely deteriorated and many skilled health workers and clinicians have left the country. All these factors have contributed to the disastrous decline in the medical health sectors in Iraq and resulted in an increase of a wide range of different tumours and diseases throughout the country (Al Hilfi et al., 2013).



Figure 1.1: The major polluted sites with DU in Iraq (Al Hilfi et al., 2013).

This research study was conducted in collaboration with the MRI Unit of Al-Kadhimiya Teaching Hospital in Iraq. This unit has faced many problems in diagnosing and issuing diagnostic reports for the massively increased number of inpatient and outpatient cases. The average number of patients received daily by the MRI unit is over 110 patients, totalling about 2640 patients scanned monthly (Hasan and Meziane, 2016; Hasan et al., 2016a).

Brain tumours are relatively less common than other neoplasms, such as those of the lung and breast, but are considered highly important because of prognostic effects and high morbidity. They require specific studies due to their complicated pathology, making them difficult to diagnose (Karkavelas and Tascos, 2011). According to the statistical report published by the Central Brain Tumour Registry of the United States (CBTRUS), brain tumours are the second leading causes of death among children and young adults (Tonarelli, 2013). There are no specific known causes for brain tumours and these are most likely wide and varied. Many risk factors have been suggested such as head injuries, hereditary syndromes, immunosuppression, ionizing radiation, cell phones, chemicals, etc. Symptoms of brain tumours include headache, nausea, vomiting, eyesight, hearing, speech problem, balance difficulties, personality changes, memory slips, loss of concentration, seizures and in extreme cases death.

Accurate visual detection and segmentation of brain tumours are essential for clinical diagnosis, predicting prognosis and treatment and beneficial for the general modelling of pathological brain topology and the exploration of the anatomical construction of the brain and any tumours it may contain (Nabizadeh, 2015; Guo et al., 2015). The generation of detailed descriptive brain tumour information can be used to index large archival databases of medical images of brain tumours which can then be used for studies and training purposes. Ultimately this information may help clinicians and radiologists to diagnose and treat current patients by determining the previous effectiveness of applied treatments and procedures with similar tumour characteristics (Saha et al., 2012; Ray et al., 2008a; Ray et al., 2008b).

Imaging studies are important to reach a diagnosis of brain tumours (Tonarelli, 2013). Diagnostic imaging has become an invaluable tool in medicine today. Typical medical imaging techniques such as ultrasonography (US), Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) and other imaging modalities have significantly increased the knowledge of anatomy and disease diagnosis in medical research and are considered a vital component in diagnosis and treatment planning (Pham et al., 1998). Among these medical technologies, MRI is the more useful and appropriate imaging technique for brain tumours. It presents detailed information about the type, position and size of the tumours in a non-invasive manner. Additionally, it is capable of differentiating soft tissue, with high resolution and more sensitive to local changes in tissue density since this reflects the physiological alternation that can be detected and visualized by MRI. Spatial resolution which represents the digitization process to assign a number to each pixel in the original image, has increased significantly in recent years and a $1 \times 1 \times 1$ mm voxel size is now achievable (Mortazavi et al., 2012; Berry, 2007). Furthermore, MRI is different from other technologies because of its ability to produce multiple images of the same tissue with different contrast visualization and different image acquisition protocols. These multiple MRI images provide additional useful anatomical information to help the clinicians to study the brain pathology more precisely. Indeed, quantitative analysis of MRI brain scans to acquire knowledge about the human brain structure has increased dramatically in recent years because of a variety of diseases that can alter the brain structure. By analysing these alternations, it is possible to understand clearly these diseases and potentially diagnose them quickly and accurately (Mortazavi et al., 2012; Nabizadeh, 2015). Most common MRI modalities are T1-weighted image (T1-w), T2-weighted image (T2-w), fluid attenuated

inversion recovery (FLAIR) and T1 weighted images with contrast enhancement (T1c-w) (Mortazavi et al., 2012; Drevelegas and Papanikolaou, 2011). Most of the brain tumours appear as hypo-intense relative to normal brain tissues on T1-w images and hyper-intense on T2-w images. Therefore, T2-w images are commonly used to determine an initial assessment, identifying tumour types and distinguishing tumours from non-tumour tissues (Tonarelli, 2013). A contrast material is commonly used to enhance the tumour boundary against the surrounding normal brain tissue on T1-w images. This enables tumour detection that cannot be distinguished and recognized from T2-w and T1-w images, because of similarity with adjacent normal brain tissue (Drevelegas and Papanikolaou, 2011; Belkic and Belkic, 2010). Figure 1.2 shows samples of T2-w, T1-w, FLAIR and T1c-w pathological slices. In clinical routine, a T2-w scan is performed immediately after patient positioning to identify the tumour location. T1-w scan is used before and after contrast administration for tumours showing contrast enhancement. The T2-w scan in axial viewing with FLAIR is used to show non-enhanced tumours (Tonarelli, 2013).

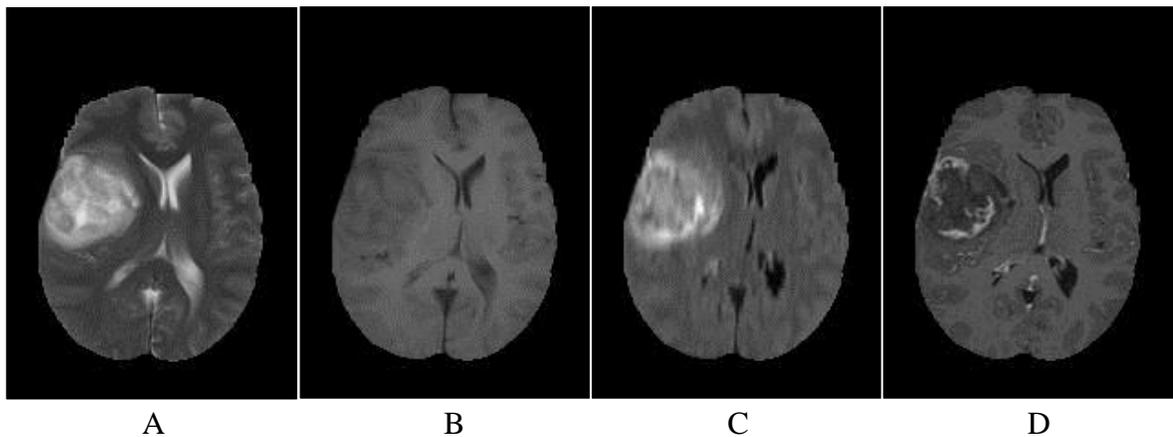


Figure 1.2: Samples of four pathological MRI slices: A) T2-w, B) T1-w, C) FLAIR and D) T1c-w.

One particular challenge in imaging features is the similarity between tumours located inside the brain white matter (WM) and those that overlap intensity distributions with the grey matter (GM). Ambiguity in classification of pixels within the tumour region can lead to inaccurate segmentation occurring when some parts of the tumour cannot be distinguished from WM/GM, due to the finite intensity resolution of the MRI image and the complexity of the human brain

anatomy. This pattern is particularly evident at the boundary between a tumour and the surrounding tissue. These boundary features are generally known as partial volumes (PV) and contain a mixture of different tissue types (Tohka, 2014). Practically, the PV affects a much wider area of the MRI image by blurring effects and mixing the intensity value of each voxel with its neighbours (Zhang et al., 2001). For instance, a 50/50 mixture of fat (hyper-intense) and bone (hypo-intense) will give a mid-range intensity value (McRobbie et al., 2007). The thicknesses of the MRI slices (5–7 mm) produce significant PV effects, in which individual slice pixels describe more than one tissue type. As a result, peripheral tumour regions are misclassified. A similar problem occurs toward the outer brain edge, where the cerebrospinal fluid (CSF) and GM overlap with the image sample. This circumstance may generate image intensities that erroneously indicate tumour presence (Mortazavi et al., 2012).

For any segmentation algorithm that has potential use in medical application and brain tumours are evaluated by either qualitative analysis, which compares the result of segmentation with reference to a standard visually (gold truth dataset or clinician evaluators) or quantitative analysis that concerns the accuracy of segmentation and how the result of segmentation is close to the reference standard by counting the number of pixels which are correctly/incorrectly identified (Berry, 2007).

In clinical routine, clinicians spend an increasing time in diagnosing and interpreting medical images due to the increased utilization of diagnostic imaging. High levels of experience are required to carry out manual and accurate delineation and classification of these medical images. As scanner resolutions improved and slices' thickness decreased, an increasing number of slices are produced. Therefore, clinicians need more time to diagnose each patient from the image set, because of the increasing magnitude of data. Coupled with the increase in inpatient numbers, this puts pressure on resources and services resulting in significant delays to both diagnosis and treatment. Therefore, automated tumour classification and segmentation have attracted considerable attention in the past two decades, resulting in many algorithms being developed for automated, semi-automated and interactive segmentation of brain tumours (Menze et al., 2015). While there has been a significant development of segmentation algorithms, they are rarely used due to wide variations of size, shape, location and feature intensity of brain tumours (Prastawa et al., 2004).

This study is concerned with developing and evaluating an efficient automated screening system for the classification and segmentation of tumours in MRI scans. It comes up with a new brain tumours classification method to discriminate normal and pathological patients based on single MRI modality and without any clinician's interception; hence reducing clinicians' examination and interpretation time. The challenge is to detect the pathological cases and identifying the most important MRI tumour slices for further detailed investigation by clinicians. The vast majority of previous studies have focused on two dimensional segmentation methods to segment the brain tumours, these are then merged to obtain a three-dimensional model (Nabizadeh and Kubat, 2015; Mikulka and Gescheidtov, 2013; Kaus et al., 2001; Prastawa et al., 2004). These methods do not exploit all the features of volumetric MRI scan and achieved lower quality of segmentation (Rousseau, 2009).

1.2 Research Questions

In order to reduce erroneous diagnostic interpretation of brain tumours in MRI scans and workload, as well as helping the clinicians to ignore the MRI brain scans of the patients who have normal brain quickly and focus on those who have pathological brain, the following research questions need to be addressed.

1. How to develop a new brain tumour detection system that classifies the MRI brain scans into normal and pathological patients more accurately?
2. Which pre-processing methods that should be used to improve the classification accuracy of brain tumours in MRI scans?
3. Which texture features can be used to classify MRI brain scans into normal and abnormal?
4. How to identify the location of brain tumours in MRI scans?
5. How to increase the segmentation accuracy of brain tumours in MRI scans?

1.3 Aim and Objectives

In attempting to answer these questions, this study aims to come out with an automated system for the classification and segmentation of brain tumours in MRI scans which will enhance the classification and segmentation accuracies. A successful system would then enable clinicians to trivially reject the MRI brain scans of healthy patients quickly and focus on those who have

high likelihood pathological brain features. This should also improve the accuracy of the diagnosis process because they will spend more time with the patients with identified abnormal MRI brain scans.

The above aim will be accomplished by fulfilling the following research objectives:

1. To investigate different brain tumour detection systems that classify the MRI brain scans into normal and pathological patients more accurately.
2. To identify which image pre-processing methods can be used to improve the classification accuracy of brain tumours in MRI scans.
3. To develop a new texture feature extraction method from MRI brain scans to improve the classification accuracy of brain tumours.
4. To develop a new automatic method for locating and identifying brain tumours in MRI scans.
5. To develop an efficient segmentation technique of brain tumours in MRI scans.

In order to answer the research questions and achieve the objectives of the research, a set of algorithms and methods are reviewed, studied and evaluated in this study.

1.4 Contribution of the Study

The main contribution of this study is to enhance the classification performance of detecting brain abnormality based on statistical texture features of MRI brain scans and the segmentation accuracy based on three-dimensional technique. Several image processing, classification, segmentation techniques that are more suitable for this domain, were selected based on previous studies as reported in the literature review. The classification and segmentation accuracies were further improved by developing novel methods.

The datasets include conventional MRI scans that were collected from Al-Kadhimiya Teaching Hospital in Iraq and BRATS 2013 which is a standard dataset was used for evaluating the classification and segmentation of brain tumours. The main contributions of this study are summarised as follows:

1. A single statistical texture feature method (MGLCM) is proposed to extract the texture features for classifying MRI brain scans into normal and abnormal with a high accuracy rate. The texture features will be extracted from a single modality of MRI (T2-w images) in an axial viewing instead of using multi-modalities of MRI (e.g. sagittal images and coronal images). The existing algorithms concentrate on combining different feature extraction methods and handling high dimensional and redundant features. As well as, utilizing multi-modalities can bring a lot of redundant information that increase the computational time and segmentation error. The study conducted comprehensive comparative studies with the existing algorithms to ensure the accuracy of the results.
2. The development of a new method (BBBGA) for identification and localisation of the abnormality in MRI brain scans, which is used to enable clinicians to look immediately at the pathological slices and avoid wasting time with investigating normal slices. Additionally, this helps to initialize the segmentation process to start automatically without the need for any human intervention or initialization.
3. The three-dimensional segmentation method (3DACWE) is used for the first time in MRI brain tumour segmentation.

1.5 Research Methodology

The work is predominately formative, being concerned with the definition of methods and concepts of the proposed system. The process of research was devised to achieve the aim and objectives of this study by determining what should be done within the system development lifecycle and how the system could be managed. The methodology that was used in this research includes the following four phases.

A. Requirement gathering and analysis

- A critical review of previous relevant works was undertaken to get a good understanding of all requirements for developing the system. These requirements are vital to identify the aim, objectives and the research problems. In addition to identify the advantages and weaknesses of existing systems by evaluating and analysing them to discover problems of previous automated systems for screening MRI brain scans.

- An MRI dataset was collected from MRI unit of Al-Kadhimiya Teaching Hospital in Iraq to evaluate the proposed system. This dataset was collected from different MRI scanners of different manufactures. In addition, BRATS 2013 dataset (Menze et al., 2015) was used to evaluate the proposed system. Both datasets include four modalities of MRI scanning; T2-w images, T1-w images, T1c-w images and FLAIR images, and are anonymous to the author and the number of patients used in each experiment is the maximum number of cases available at the time of the study.

B. System design

- All the necessary requirements to develop a system were collected, such as the requirements for pre-processing, classification and segmentation of MRI brain scans. This phase includes smoothing MRI slices in order to reduce motion artefacts and field inhomogeneity, background of MRI slices elimination, intensity standardization in order to deal with MRI slices from different MRI scanners. Finally, correcting the MSP of the brain if there is a skewness in the patient head.
- The textural features extraction that could be used to detect the abnormality of MRI brain scan by measuring the symmetry between the left and right hemispheres of the brain. These features were used to discriminant the normality and abnormality of MRI brain scans. Features preparation included feature selection and normalization.
- Two methods were used to prepare the extracted features for classification to improve the classification accuracy.
- Three classifiers were used to distinguish and differentiate brain tumours in MRI scans. These classifiers are linear discriminant analysis (LDA), support vector machine (SVM) and MLP.
- A novel brain tumour locating method was used to search and identify the location of the most dissimilar regions between the left and right hemispheres of the brain automatically without the need for user interaction.

- A fully automated brain tumour segmentation method independent of atlas registration was applied on T1-w, T2-w, T1c-w and FLAIR images.

C. Implementation

- MATLAB R2013a software was used to implement the proposed system.

D. Evaluation

- Evaluate the performance of the proposed system to ensure that the aim and objectives are achieved by using qualitative and quantitative measures.

Figure 1.3 illustrates how our research was designed and distributed over the four phases.

1.6 Thesis Organization

The thesis is organised into nine chapters describing the various phases of the research development process. These are summarised as follows:

Chapter 1: This chapter provides a brief introduction to the research described in this study. It includes a brief introduction to the clinical background to diagnosis using MRI scans, which leads to the important task of identifying brain normalities and abnormalities. The chapter states the research questions and the aim and objectives of the research. The contribution of the study, research methodology and a brief description of the remaining chapters of the thesis complete the content of the first chapter.

Chapter 2: This chapter gives a comprehensive survey of previous works related to the automated detection and segmentation of brain tumours in MRI scans. The full analysis of the techniques and methods used are also described. In addition to providing some background and investigations into some image processing techniques such as image enhancement techniques, intensity normalization, features extraction techniques, features transformation techniques, classification techniques and image segmentation techniques.

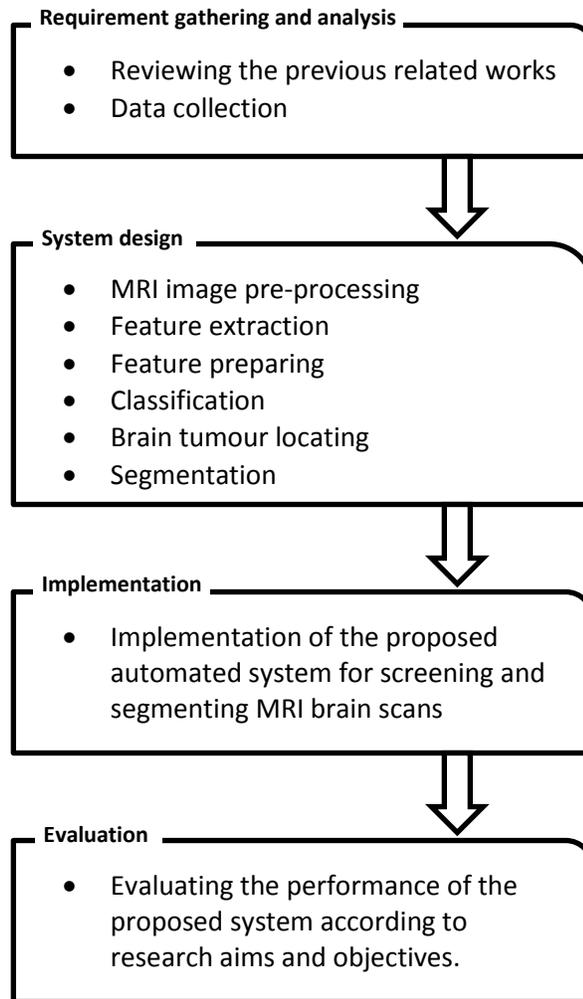


Figure 1.3: The research phases of the proposed system.

Chapter 3: This chapter presents a brief introduction to digital image representation, followed by describing the important characteristics of MRI. In addition, a brief description of Al-Kadhimiya Teaching Hospital and brain tumours are included.

Chapter 4: This chapter presents the pre-processing algorithms that are primarily implemented on MRI brain scanning images. These algorithms include resizing MRI slices dimensions, MRI enhancement algorithms, intensity normalization, background elimination and MSP detection and correction.

Chapter 5: This chapter describes in details the proposed method for texture feature extraction by MGLCM. It also includes the implementation of the stepwise Analysis of Variance (ANOVA) based feature selection method. The ANOVA is assessed by comparing it with other methods by observing the classification performance of used classifiers. This chapter presents a comparison between the outcomes of three classifiers for classifying MRI brain scans. These classifiers are LDA, SVM and MLP as an intelligent classifier.

Chapter 6: This chapter presents and describes the proposed method for texture feature extraction by three-dimensional modified grey level co-occurrence matrix (3DMGLCM). It includes an implementation of the stepwise ANOVA based feature selection method. It also includes a comparison between the performance of MGLCM and 3DMGLCM.

Chapter 7: This chapter presents the detailed design of the proposed BBBGA method. This method is used to locate and identify the location of brain tumour in axial viewing of MRI scan. The implementation of this method is described in details.

Chapter 8: This chapter presents experiments conducted to segment brain tumours automatically by using 3DACWE after being initialized by BBBGA. A summary of all the experiments is given and compared with the two-dimensional active contour without edge (2DACWE) to evaluate the effectiveness of 3DACWE.

Chapter 9: This chapter reviews the proposed system that has been done and the techniques that have been used in this study. The chapter concludes by discussing the system, issues and results achieved in this research followed by suggestions for future works and improvements.

CHAPTER TWO

Literature Review

Overview

This chapter presents a survey of previous works and the theoretical background of three main topics; texture feature extraction, classification and segmentation. The best techniques will be identified to justify their use in the current research.

2.1 Introduction

The development of medical imaging over the last four decades has revolutionised medical diagnosis and it is widely used in many procedures that include three-dimensional volumetric visualization of CT and MRI data of the spine, internal organs and brain. The medical imaging techniques are now able to investigate the structure, function and pathology of the human body with a variety of imaging systems and used to plan treatment and surgery (Brody, 2009). Medical imaging is a discipline within the medical field which involves the use of technologies to produce images of the internal structures of the human body in a way which is as non-invasive as possible (Birry, 2013). There is a variety of medical imaging technologies which are used to help the clinicians to identify pathological conditions inside the body, congenital defects, functionality of the organs and vessels, broken bones and tumours. Due to the increasing number of medical imaging technologies, the use of computers in facilitating their processing and analysis has become essential. Practically, computer algorithms for the delineation of anatomical and other regions of interest are a key component in assisting and automating specific radiological tasks such as the detection and classification of tumours, construction of grey-scale or colour histogram, segmentation, area measurement, etc. (Pham et al., 1998, Birry, 2013).

The computer applications that support medical imaging techniques, are using image processing algorithms for quantitative analysis to help clinicians who are currently assessing and diagnosing medical images visually, which has some limitations in terms of time and accuracy.

The reasons behind these limitations are inter-observer variations and error due to stress, oversight and limited experience. The observer usually hangs and reads films on the alternators and the inspection of the scans is prone to errors owing to visual exhaustion after spending long hours of reading. The diagnosis of medical images by experts is known to be a subjective assessment (Liu, 2009). Hence, computer analysis can be used to add more objectivity to the subjective diagnosis. Therefore, computer analysis becomes essential in improving diagnostic accuracy and confidence even for experts with high experience. The imaging of human organs using MRI has been the subject of many research projects including the detection of different types of tumours such as breast, lung, liver and brain. Of particular interest to this research, is the reduction of the time spent by the clinicians on normal cases to allow them to spend more time on abnormal ones.

The need to develop an automated screening and segmentation system of brain tumours has increased rapidly over the last decade. A general framework of screening system includes dataset collection, pre-processing, features extraction, feature selection, classification and brain tumour segmentation as illustrated in Fig. 2.1. In order to choose the most suitable methods for each step in the proposed system, a thorough review of the research undertaken in the various phases is reviewed and used as the justification of the choices we have made in this work.

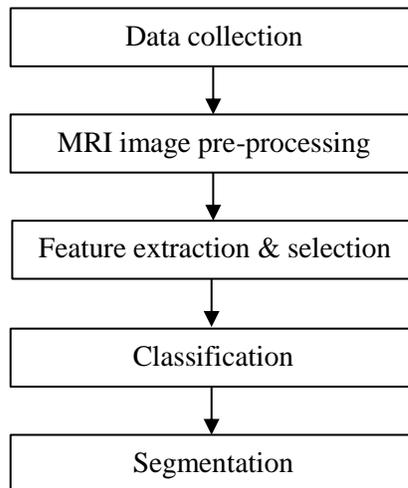


Figure 2.1: A general framework of the screening system.

2.2 Pre-processing Analysis

Pre-processing analysis is an important step in the proposed systems. Typically, it includes image enhancement to reduce the effects of corruptions that could contaminate medical images during the acquisition or transmission process (Nowak, 1999; Aelterman et al., 2008; Lauwers et al., 2009; Anju et al., 2013; Bovik, 2009).

2.2.1 Image Enhancement

Image enhancement techniques are widely used to refine medical images in order to improve the visibility of important structures and assess the visual information. The desired features become easier to perceive for the human visual system or more likely to be detected by automated image analysis systems. The goal of image enhancement is simply that the enhanced image should be more appropriate than the original image for the required task (William, 2001; Solomon and Breckon, 2011). Image enhancement depends strongly on the details and the specific information of medical images that users are attempting to extract but that are not visible to them. For instance, if there is an automated image analysis system that traces the outline of the edges and measures the shape and size of the outline, the image enhancement system would enhance the edge outline of the objects in the medical image (William, 2001).

Image enhancement techniques can be divided into two categories; spatial domain and frequency domain. The image enhancement in spatial domain includes convolution operation, that uses specific masks or kernels to produce smoothing or sharpening of an image (El-Shenawy, 2013; Dougherty, 2009). While the image enhancement in frequency domain is performed after transforming the images into frequency domain using transformation methods. Generally, the image enhancement in spatial domain is efficient computationally and require less processing resources to implement (Gonzalez and Woods, 2002; Birry, 2013).

The MRI data is probably affected by several sources of quality deterioration such as scanning times, movement of patients, motion of molecules in the scanning subject, respiration motion, heartbeat and the acquisition process. The noise may be propagated in MRI scans and it is required to remove it carefully in the pre-processing step. The typical and common smoothing

approaches will be used in this study for reducing noise in MRI brain scans (Nabizadeh and Kubat, 2015; Tantisatirapong, 2015; Birry, 2013; Verma et al., 2008).

2.2.2 MRI Intensity Normalization

MRI technology has become an efficient tool for the diagnosing of brain diseases. However, there is an obstacle with quantitative texture analysis of MRI images due to intra-scan and inter-scan image intensity variations between the same and consecutive MRI slices due to MRI scanners. Thus, the extracted results are not comparable between consecutive or repeated scans or within the same scan, between different anatomic regions (Loizou et al., 2009; Nabizadeh, 2015). In addition, acquiring MRI data from different scanners at different sites produces variance in the dynamic intensity range of the brain tissue even though they are used identical acquisition protocol. Furthermore, the variation in the intensity of MRI brain scans can vary significantly due to different manufacturers and scanner-models, bias field and different pulse sequence parameters (Tantisatirapong, 2015). Therefore, image intensity of MR slices should be normalized and standardized and this may have a significant impact on the clinical diagnosis, image analysis and computer aided diagnosis. There are six MRI intensity normalization techniques; contrast stretch normalization, intensity scaling, histogram stretching, histogram normalization, Gaussian kernel normalization and histogram equalization. In this study, the histogram normalization method will be used in the pre-processing step before texture feature extraction due to its performance compared to other normalization techniques (Loizou et al., 2009; Nixon and Aguado, 2008). It stretches and shifts the original MRI slice histogram in order to cover all the grey scale levels in the MRI slice as given in Eq. 2.1 (Loizou et al., 2009; Watt and Policarpo, 1998; Tantisatirapong, 2015; Nabizadeh and Kubat, 2015).

$$f(x, y) = \frac{G_{Higher} - G_{Lower}}{(G_{max} - G_{min})} (G(x, y) - G_{min}) + G_{Lower} \quad 2.1$$

where $G(x, y)$ is the original MRI slice, starting at a minimum grey level value G_{min} and extending to G_{max} , $f(x, y)$ is the normalized MRI slice within a minimum grey level G_{Lower} , and a maximum grey level G_{Higher} . x and y are the coordinates of pixels in MRI slice.

2.2.3 Mid-Sagittal Plane of Brain Detection

The Mid-Sagittal Plane (MSP) of the brain is a plane that separates it into two halves known as the two hemispheres of the brain. Identifying this plane is considered important for many automated systems that measure the similarity between the two hemispheres. Therefore, the detection of MSP is a topic that has been investigated for decades (Kuijf et al., 2014). Liu et al. (1998), Liu and Collins (1996) and Ardekani et al. (1997) proposed an automated algorithm for detecting the MSP based on the symmetry axis that should have the same orientation of the patient's head. Hence, the process is based on searching for the orientation of the reflection line that maximizes the cross-correlation between the original image and the rotated image. Bergo et al. (2008) proposed an automated method for detecting the longitudinal fissure, which is clearly visible in T1-w images. The author assumed that the MSP contains a maximal area of CSF, which appeared as a low intensity area. Therefore, the proposed method was based on searching for a sagittal plane that minimized the intensity mean. Ruppert et al. (2011) proposed an algorithm for extracting the MSP by searching the plane that maximizes a bilateral symmetry measure. The bilateral symmetry measurement was based on extracting the edge features from the MRI brain slice. Then measuring the similarity using the correlation between the left and right hemispheres with respect to a candidate cutting plane. Jayasuriya and Liew (2012) proposed an automated algorithm for detecting the MSP of the brain by exploiting the property that the longitudinal fissure in T1-w images appears as a dark area. A set of lines were drawn in multiple angles to analyse the intensity along these lines. The best possible line that fits the inter-hemispheric fissure which represents the angle of the MSP to the vertical axis was chosen. Nabizadeh and Kubat (2015), Ray et al. (2008a) and Saha et al. (2012) separated the brain into two hemispheres by finding the longest diameter that represents the MSP of the brain. Their algorithms included separating the brain from the background, finding the brain centre, finding the brain's borderline, determining the lengths of all possible brain diameters and assigning the longest diameter as the MSP of the brain.

Previous works have used different techniques for detecting MSP of the brain as summarized in Table 2.1. However, the intensity-based analysis methods (symmetry, fissure) might not be the optimal solution to identify the MSP, because they are sensitive to any pathological conditions that could induce asymmetries and displacement of anatomical structures of the brain

(Kuijf et al., 2014; Liu et al., 1998; Liu and Collins, 1996). In this study, the emphasis is on estimating the orientation of the skull that is identical to the reflection line and passes through the MSP of the brain (Hu and Nowinski, 2003; Liu and Collins, 1996).

Table 2.1: Summary of existing MSP methods.

Method	Features	Measure
Junck et al. (1990)	Symmetry	Intensity cross correlation
Liu and Collins (1996)	Symmetry	Intensity cross correlation
Ardekani et al. (1997)	Symmetry	Intensity cross correlation
Liu (2009)	Symmetry	Intensity cross correlation
Prima and Ourselin (2002)	Symmetry	Intensity cross correlation
Tuzikov et al. (2003)	Symmetry	Intensity cross correlation
Hu and Nowinski (2003)	Symmetry	Local symmetry index
Ruppert et al. (2011)	Symmetry	Correlation
Jayasuriya and Liew (2012)	Fissure	Minimized the intensity mean
Bergo et al. (2008)	Fissure	Minimized the intensity mean
Ray et al. (2008a)	Longest diameter	Lengths of all possible brain diameters
Saha et al. (2012)	Longest diameter	Lengths of all possible brain diameters
Nabizadeh and Kubat (2015)	Longest diameter	Lengths of all possible brain diameters

2.3 Texture Analysis

The texture is a variation of the data at scales smaller than the scales of interest. It represents an intrinsic property of the imaged object and becomes a valuable cue in relation to the object classification and it is essential to describe the texture in an objective way, independent of human perception and visual abilities (Petrou, 2011). The fundamental objective of any diagnostic imaging investigation is tissue characterization; therefore, the texture analysis is considered to be an efficient way to extract higher-level information. This information could be used to distinguish a primitive characteristic or attribute of medical images (Nabizadeh and Kubat, 2015; William, 2001). Texture analysis is one of the image processing fields that is still a challenging problem in computer vision (Jähne, 2005). It is a potentially valuable and versatile tool in neuro-MR imaging and provides an alternative diagnostic tool for MR image analysis (Kassner and Thornhill, 2010). It has been used to assess MRI slices of biological tissues, which contain large amounts of microscopic details that are scarcely addressed by visual inspection. Moreover, texture analysis can characterize patterns of tissues in MR slices better than the human visual system because it is more sensitive to variations of grey-level intensity in medical images. However, texture analysis was proved to produce better discrimination between healthy

and pathological tissues compared to human visual examination (Tantisatirapong, 2015). It is more robust for monitoring disease progression or treatment response with time (Kassner and Thornhill, 2010).

Texture features techniques are classified according to their domain and are categorized into spatial texture features and spectral texture features. The spatial techniques are used to analyse the spatial distribution of grey values by computing local features at each point in the image, and deriving a set of statistics from the distributions of the local features such as convolution filters, co-occurrence matrix and spatial autocorrelation. While the spectral techniques transform an image to frequency domain and then texture features are determined from the transformed image such as Fourier transform, wavelet transform and cosine transform (Tian, 2013; Haralick et al., 1973). However, the spatial techniques have been more prevalent and advantageous than spectral techniques because the spatial distribution of grey values provides the defining qualities of texture (Nabizadeh, 2015).

There are several studies that used statistical features in discriminating objects in images and sometimes they are combined with other features extraction techniques. Liu et al. (2012) proposed a hybrid method by combining two-dimensional discrete wavelet transform (2D DWT) and Gabor wavelet for texture feature extraction which can be used to recognize different categories of brain tumours. Where, the Gabor wavelet was used with eight orientations and various frequencies, while the 2D DWT was used for noise removing as a prior pre-processing step. LDA was used to evaluate these features based on k-fold cross-validation method. Similarly, Lahmiri and Boukadoum (2013) developed a new methodology for automatic features extracting from biomedical images using 2D DWT and Gabor wavelet with different frequencies and spatial orientations. The classification was performed using SVM and accuracies of 86%, 68% and 50% were achieved on MRI brain images, mammograms and retina respectively. Kharrat et al. (2010) and Beura et al. (2015) used grey level co-occurrence matrix (GLCM) and 2D DWT to extract texture features in their studies. Kharrat et al. (2010) presented work that classified brain tumours in MRI brain scans into normal, benign and malignant. The proposed system worked on T2-w images in axial viewing. Generally, in medical image analysis, the determination and classification of tissue type (normal or abnormal) are performed using texture features, where MRI image texture is sufficient to be used to determine the tumour

type. The texture features were extracted by using GLCM from low frequency band after decomposing the MRI image using 2D DWT into two levels. The genetic algorithm (GA) was used to extract the most relevant texture features and the classification was done using SVM. 97% accuracy was achieved in classifying a dataset of 83 MRI brain slices. Beura et al. (2015) used these texture features to classify the breast tissues into normal, benign and malignant tumours by using mammogram images. Then the most relevant features were selected using the F-statistic method. The classification was performed using the back propagation neural network (BPNN). As well as, 2D DWT was also used for feature extraction from MRI brain scans in (Saritha et al., 2013; Kalbkhani et al., 2013). However, 2D DWT has some limitations in capturing relevant information, and a lack of translation-invariant where any simple shift in the image results to significant modifications in values of wavelet coefficients (Baaziz et al., 2010; Du et al., 2016; Tantisatirapong, 2015). While, Gabor wavelet has some drawbacks such as computation complexity, production a huge number of redundant features at different scales and difficult to localise a spatial structure of natural textures (Tantisatirapong, 2015; Baaziz et al., 2010). Pantelis (2010) developed a medical system to classify and discriminate the normality and abnormality of MRI brain slices by combining three approaches for texture features; GLCM, first order statistical method and grey level run length matrix (GLRLM). Additionally, the dimensionality of the extracted features was reduced using the Wilcoxon test method which is a non-parametric statistical hypothesis test method. The most relevant features were retained when the P-value is less than 0.001, and SVM was used to classify a dataset of 67 patients and the maximum classification accuracy obtained was 93%. First order statistical method was also used in (Bauer et al., 2011). However, the power of first order statistical method for distinguishing between textures with different spatial arrangement and having the same grey value distribution is limited. Thus many textures cannot be distinguished using first order statistical features (Pantelis, 2010; Jähne, 2005; William, 2001; Padma and Sukanesh, 2011; Nailon, 2010; Kassner and Thornhill, 2010). As well as, GLRLM is inefficient compared with other feature extraction techniques such as GLCM (Tantisatirapong, 2015). Where, the extracted features by GLRLM do not include maximum texture information and many of these features are highly correlated with each other (Tang, 1998). Gomez et al. (2012) proposed an automated system to classify breast lesions in ultrasound images using GLCM method. LDA was used to classify the extracted features and the maximum acceptable classification rate was

87%. Qinggang et al. (2015) combined in their algorithm two texture features extraction methods; Gabor wavelet and GLCM methods in addition to principal component analysis (PCA) that was used to optimize the extracted features. Similarly, Sachdeva et al. (2013) proposed a multiclass brain tumour classification algorithm using various techniques for feature extraction; Laplacian of Gaussian (LoG), GLCM, rotation invariant local binary patterns (RILBP), intensity-based features (IBF) and Gabor wavelet. Finally, these features were classified using an artificial neural network (ANN) after using PCA for data reduction. An overall classification accuracy of 91% was achieved to classify a dataset of 428 MRI brain scans. However, the main drawback of local binary pattern (LBP) is that the spatial relations among LBPs are mostly discarded within the LBP histogram generation process, because they are picked into a single histogram and leads to a loss of global image information (Mohammadi et al., 2012). Nabizadeh and Kubat (2015) proposed a fully automated algorithm using five effective texture-based statistical feature extraction methods; first order statistical features, GLCM, GLRLM, histogram of oriented gradient (HOG) and LBP. PCA was used for feature dimension reduction and 97.4% accuracy was achieved for classifying the brain scans of 25 pathological patients using SVM. Hackmack et al. (2012) proposed an approach for analysing MRI brain slices for diseases classification. Dual-tree complex wavelet transform was used for feature extraction. The classification was performed using SVM to classify the extracted features into normal and abnormal MRI brain scan. Ain et al. (2010) used discrete cosine transform (DCT) method for feature extraction to develop an automated system to classify the MRI brain slices into normal and abnormal. A Bayesian classifier was used as a statistical classifier in this study. Sachdeva et al. (2016) developed a system for assisting radiologists to classify brain tumours. The proposed system is composed of three main modules; first brain tumours were segmented using a semi-automatic content based active contour model (CBAC). Second features extraction using GLCM, LoG, Gabor wavelet, RILBP, IBF and shape based feature (SBF) were performed. Then it was followed by features selection to measure the significance of features using GA. Finally, SVM and MLP were used independently to classify brain tumours and a comparison of their performance was conducted. The accuracies achieved by SVM and MLP were 91.7% and 94.9% respectively.

The previous works showed different algorithms and techniques that were used for features extraction from MRI scans (e.g. brain and breast). The detection of brain tumours is generally

a more complex task than the detection of any other objects in image processing. Pattern recognition algorithms usually depend on the shape of the objects or colour information while brain tumours have irregular shapes and no colour information that could be used to discriminate the pathological brain. Another widely used knowledge about brain tumours, is the symmetry features between the hemispheres of a healthy brain (Dvořák et al., 2013; Saha et al., 2012; Ray et al., 2008b). This study is primarily concerned with adapting texture feature from single MRI modality in order to measure the degree of symmetry between the two hemispheres of the brain to discriminate abnormality. The second order statistical features techniques (e.g. GLCM) will be considered because it is still the most popular approach for deriving the most popular spatial statistical features for texture analysis of biomedical images and provide information about the spatial arrangement and intensities distribution in MRI slices. It has a good performance that outperforms other techniques such as wavelet features, Fourier features and Gabor wavelet (Materka and Strzelecki, 1998; Kharrat et al., 2010).

In the following subsections, we review some of the approaches and techniques used for texture analysis as used in previous researches on MR brain scans.

2.3.1 Grey Level Co-occurrence Matrix

The grey level co-occurrence matrix is one of the most popular statistical techniques for extracting second order statistical texture features from grey-level images and estimating the relationship among pixels or groups of pixels. It considers the spatial relationship between pairs of pixels by determining the occurrence appearing in the image (Haralick et al., 1973; William, 2001; Nabizadeh, 2015; Nailon, 2010). The GLCM considers the relative distance between pixels instead of their position. Consequently, unlimited GLCMs can be constructed (Petrou, 2011). It is still the most popular approach today due to its good performance and the extracted features carry information about the distribution of intensities relative to the position of pixels (William, 2001; Nixon and Aguado, 2008; Kassner and Thornhill, 2010). The GLCM has rows and columns that are equal to the number of grey levels in the image and includes information about the number of pairs of intensity value of pixels at different offset distances d in four different orientations ($\theta=0^\circ, 45^\circ, 90^\circ$ and 135°) as illustrated in Fig. 2.2.

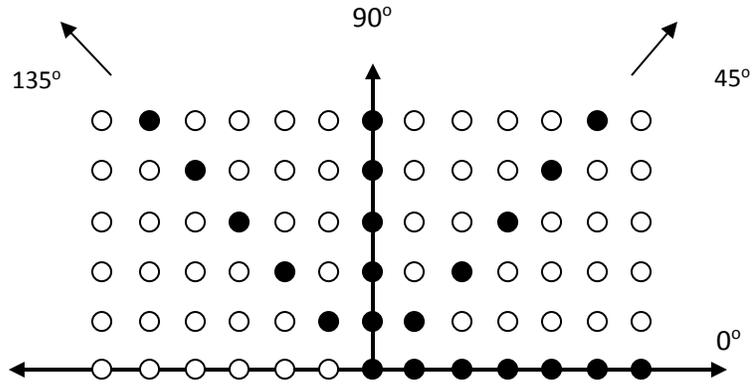


Figure 2.2: Representation of GLCM.

For an image I of size $(N \times N)$, the GLCM is calculated using Eq. 2.2.

$$GLCM_{\theta}(i, j) = \sum_{x=1}^N \sum_{y=1}^N \begin{cases} 1, & \text{if } I(x, y) = i \text{ and } I(x + \Delta_x, y + \Delta_y) = j \\ 0, & \text{otherwise} \end{cases} \quad 2.2$$

where Δ_x and Δ_y are the offset distances between the reference pixel of coordinates x, y , and its neighbours. i and j are the coordinates of GLCM (Eleyan and Demirel, 2009).

The 19 co-occurrence statistics that are used in this study, are derived from each GLCM and they will be explained in details in chapter five and are listed in Table 2.2 (Haralick et al., 1973; Tantisatirapong, 2015; Gomez et al., 2012; Pantelis, 2010; Zulpe and Pawar, 2012; Gebejes and Huertas, 2013; Albregtsen, 2008; Yang et al., 2012; Sonka et al., 2014; Wilson and Ritter, 2000; Nabizadeh and Kubat, 2015; Qinggang et al., 2015).

Table 2.2: List of the 19 co-occurrence statistics extracted from GLCM.

Texture Features	Texture Features
Contrast	Inverse Difference Moment Normalized
Correlation	Sum Average
Entropy	Sum Entropy
Energy	Sum Variance
Homogeneity	Difference Entropy
Dissimilarity	Information Measure of Correlation I
Sum of Square Variance	Information Measure of Correlation II
Cluster Shade	Autocorrelation
Cluster Prominence	Maximum Probability
Inverse Difference Normalized	

Although, the GLCM is an efficient technique for extracting spatial texture features from grey scale images, it requires more computational time as more levels are included in the GLCM (Tantisatirapong, 2015; Gomez et al., 2012). The example given in Fig. 2.3, shows how the GLCM can be determined for a given matrix with four orientations ($\theta=0^\circ, 45^\circ, 90^\circ$ and 135°) and an offset distance of 1.

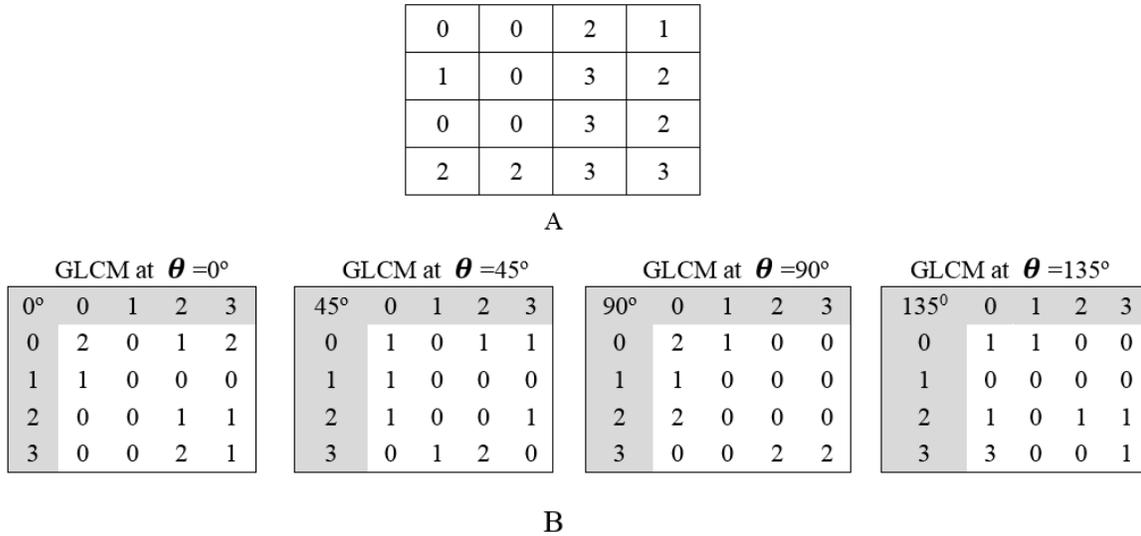


Figure 2.3: Implementation of GLCM; A) Original Matrix, B) The four GLCM with $d=1$ and orientations ($\theta = 0^\circ, 45^\circ, 90^\circ$ and 135°).

2.3.2 Gabor Wavelet

Gabor wavelet is a frequency transform method and one of the most popular signal processing based texture features extraction method. It was proposed to be used to model the responses of the human visual system. It includes a bank of filters with different spatial frequencies and orientations. It encodes the texture features of an image into multiple narrow corresponding spatial frequency and orientation channels (Howarth and Ruger, 2005; Nabizadeh and Kubat, 2015). In the spatial domain, a two-dimensional Gabor function is a Gaussian modulated complex sinusoidal function, the complex exponential has a spatial central frequency f and an orientation θ . A two dimensional Gabor function can be defined as Eq. 2.3 (Howarth and Ruger, 2005; Qinggang et al., 2015; Kong et al., 2003; Nabizadeh, 2015).

$$G(x, y) = \frac{f^2}{\pi\gamma} \left(-\frac{(x'^2 + \gamma^2 y'^2)}{2\sigma^2} \right) \exp(j2\pi f x' + \phi) \quad 2.3$$

where x' and y' are determined by using Eq. 2.4 and Eq. 2.5.

$$x' = (x \cos \theta + y \sin \theta) \quad 2.4$$

$$y' = (-x \sin \theta + y \cos \theta) \quad 2.5$$

where σ is the width of the Gaussian envelope, γ is the spatial aspect ratio which specifies the ellipticity of the support of the Gabor function, and ϕ is the phase offset (Nabizadeh, 2015; Sachdeva et al., 2013).

Image texture features g can be extracted by convolving the MRI image I with Gabor wavelet G at specific frequency f and orientation θ , as given by Eq. 2.6 (Liu et al., 2012).

$$G(x, y, \sigma, f, \theta) = I * g(x, y, \sigma, f, \theta) \quad 2.6$$

The majority of researchers (Nabizadeh and Kubat, 2015; Liu et al., 2012; Lahmiri and Boukadoum, 2013) and in this work use Gabor wavelet in five scales and eight orientations, meaning that there are forty Gabor wavelets. Since the adjacent pixels in the MRI image are highly correlated, there are many redundant features that are produced by Gabor wavelet. Consequently, dimensionality reduction methods are required to reduce the size of feature vectors (Kharrat et al., 2010; Haghghat et al., 2015).

2.4 Preparing the Extracted Features for Classification

There are some pre-processing steps that should be performed to prepare the extracted features for classification to improve the accuracy, performance, efficiency and scalability of the classification process (Han et al., 2011).

2.4.1 Relevance Analysis

It is a process of removing irrelevant and redundant features by finding a minimum set of features such that the resultant probability distribution of the data classes is as close as possible to the original distribution of features (Han et al., 2011; Tang et al., 2014). High dimensional

datasets usually include hundreds of features and some of these features are irrelevant or redundant and may lead to deteriorating the classification performance. The irrelevant feature can be removed without any effect on distinguishing classes or the classification performance. While, some of these features behave in a similar way and then simultaneous presence as redundant features. Therefore, removing one of them will not affect the classification performance (Gomez et al., 2012). It is expected that the result of relevance analysis is completely correlated with the results of the classification (Stańczyk and Jain, 2015). Most machine learning algorithms can suffer from both insufficient and excessive number of features and they are designed to use the most appropriate features for making their decisions. The most appropriate features denote the most promising features that are used to split the given data into classes with more discriminating power (Witten et al., 2011). Therefore the relevance analysis becomes most popular and widely used in pattern recognition, data analysis, multimedia information retrieval, medical data processing, machine learning and data mining applications (Tu et al., 2007; Stańczyk and Jain, 2015). The relevance analysis makes features easier to understand, reduces complexity and computational cost, enhances interpretability of feature, reduces storage requirements, reduces training and utilization times and improving the classifier performance by reducing misclassified data and generalization error (Nabizadeh, 2015; Han et al., 2011; Tantisatirapong, 2015; Pantelis, 2010). There are two methods that are used to perform the relevance analysis: correlation analysis and features selection. Correlation analysis is to recognize whether any two given features are related statistically. Once they are strongly correlated, one of the features could be discarded from further analysis; "a good feature subset is one that contains features highly correlated with the class and uncorrelated with each other" (Hall, 1999). Feature selection methods are used to remove the irrelevant features (Tantisatirapong, 2015), and are divided into two categories; the first is to make features assessment based on general characteristics of the data and these are called the filter method because the features are filtered to produce the most promising subset. The second category evaluates features using machine learning algorithm and they are called the wrapper methods because the learning algorithm is wrapped into the selection procedure (Witten et al., 2011; Pantelis, 2010). There are many techniques that are used for relevance analysis such as ANOVA (Johnson and Synovec, 2002; Baboo and Sasikala, 2010), PCA (Petrov and Jordanov, 2011), GA (Ortiz et al., 2013) and kernel PCA (Moghaddasi et al., 2014; Wang, 2012).

2.4.1.1 One-Way Analysis of Variance

The one-way analysis of variance (ANOVA) is a robust statistical technique that is used for data analysis for assessing the discriminatory power of each sample in the features vector based on testing whether the means of multiple groups are significantly different. The null hypothesis in ANOVA is always that there is no difference in the means of groups. The variance based on within-group variability should be equal to the variance based on between-group variability. It assumes that all the samples are normally distributed with equal variance and all samples are mutually independent (Dubitzky et al., 2007).

The ANOVA predicts the significance of feature using F-statistic and P-value. Where, the F-statistic is defined as a ratio of between-group variance to the within-group variance that is used to assess whether the ratio of these variance estimates is significantly greater than 1. While P-value is the probability of the test statistic being at least equal to or less than the critical value of the test (5% or 1%) (Dubitzky et al., 2007). When applying "the ANOVA on a two-class scenario, it is equivalent to the two-sample *t*-test assuming equal variances" (Dubitzky et al., 2007; Beura et al., 2015; Crawley, 2012).

The between-group variance is calculated by Eq. 2.7 (Johnson and Synovec, 2002).

$$MSS_B = \frac{\sum_i^n (M_i - \bar{M})^2 n}{k - 1} \quad 2.7$$

where M and \bar{M} denote the predictor and mean of predictors within each group respectively, k denotes the number of groups and n denotes number of predictors for each group.

The within-group variance is calculated using Eq. 2.8:

$$MSS_w = \frac{\sum_i \sum_j (M_{ij} - \bar{M}_j)^2}{N - k} \quad 2.8$$

where M_{ij} is the predictor of the j^{th} class and N is the total number of predictors of all classes.

Then, the F-statistic is a ratio between the two variances, as given by Eq. 2.9.

$$F - statistic = \frac{MSS_B}{MSS_w} \quad 2.9$$

From the ANOVA table, when the P-value for the F-statistic value is less than the critical value α , then the feature will be significant. The critical value α is proposed and fixed by Fisher in 1956 (Quinn and Keough, 2002). He suggested the idea of a conventional probability for accepting or rejecting a hypothesis and it was one in twenty (0.05 or 5%). The probability of the F-statistic rising from multiple groups' distributions give us a measure of the significance of the between group variation as compared to the within group variation (Baboo and Sasikala, 2010). Consequently, the predictor will be significant when $P < 0.05$, very significant when $P < 0.01$ and highly significant when $P < 0.001$ (Quinn and Keough, 2002). In order to achieve more efficient feature by ANOVA, it is necessary that the number of features is greater than the number of samples or patients within the dataset (Dubitzky et al., 2007). In this study, we will use ANOVA to analyse and measure the relevance of the extracted texture features. Because the P-value does not indicate actually to what degree each group is separated from the others and ignores the redundancy of features (Johnson and Synovec, 2002; Chen et al., 2005). This drawback is overcome by using the F-statistic to determine the power of discrimination of the features by thresholding. Where, different threshold values will be taken to ignore the redundant features and evaluate the selected features at each time by observing the performance of the classifier.

2.4.1.2 Principle Component Analysis

The principle component analysis (PCA) is a powerful linear transformation technique that has been widely used for different purposes such as dimensionality reduction, finding the most relevant variables, orientation detection, face recognition and image compression (Smith, 2002; Moghaddasi et al., 2014). It is an unsupervised feature selection technique that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables, called principal components (PCs) (Petrov and Jordanov, 2011). These PCs are a linear combination of the original features with different coefficients associated to each original features, and they are orthogonal to each other to maintain most of the variability of the features (Tantisatirapong,

2015). The PCA normalizes the input feature set to zero mean by subtracting the mean from each feature to ensure that features with a large scale will not dominate features with smaller one. The PCs are unit vectors and each PC points in the direction of a new axis and is perpendicular to the others. The desirable axis (PC_1) that has the highest eigenvalue and the most variance among the features set. The second orthogonal axis (PC_2) has as much of the remaining variance in the feature set and so on. The PCs that have low variance, can be ignored and only the strongest ones are retained which can be used to obtain a good approximation of the original data (Han et al., 2011; Wallisch et al., 2014; Sonka et al., 2014).

Let us assume that there are N PCs as the output of the PCA, and M is the number of selected PCs corresponding to the highest eigenvalues. The optimal number of selected feature M depends on the reconstruction ratio γ . Where γ is defined as the ratio of the summation of the M selected eigenvalues to the total sum of all eigenvalues and computed by Eq. 2.10 (Nabizadeh, 2015).

$$\gamma = \frac{\sum_{i=1}^M \gamma_i}{\sum_{i=1}^N \gamma_i} \quad 2.10$$

where γ_i is the eigenvalues that are sorted in a descending order, and γ is the reconstruction ratio. The number of the selected feature will increase when γ values approach 1.

PCA is computationally inexpensive and can be applied to sparse and skewed data but its main drawback is that it is not appropriate for more than two-dimensional data because the covariance matrix is proportional to the dimensionality of the data. As a result, the computation of the eigenvectors might be infeasible (Han et al., 2011; Van der Maaten, 2007). In addition, if the data has more complicated structures which cannot be represented in a linear space, the use of the traditional PCA will not be appropriate and helpful (Wang, 2012).

2.4.2 Features Transformation

Features transformation is a process of transforming and scaling the extracted features into new predetermined ranges due to the different dynamic range of the feature across all cases in the dataset (Tantisatirapong, 2015; Han et al., 2011). Therefore, the extracted features should be normalized to standardize the scale of the effect of each feature to be more appropriate for the

classification process. In addition, to avoid features with initially large numeric ranges dominating the features with initially smaller numeric ranges, and also avoid numerical complexities during the learning phase of the classification process (Kalbkhani et al., 2013). There are many normalization techniques and some of them are described in the following subsections.

2.4.2.1 Min-Max Normalization

Min-Max normalization performs a linear transformation on the extracted features with preserving the relationships among the feature vector values. The feature values are scaled from the predetermined range to the interval range between 0 and 1, as defined in Eq. 2.11 (Tantisatirapong, 2015; Han et al., 2011; Larose, 2005; Kalbkhani et al., 2013; Zhang et al., 1998; Jayalakshmi and Santhakumaran, 2011).

$$v' = \frac{v - \min}{\max - \min} \quad 2.11$$

where v' is the normalized value of feature vector v , \min is the minimum value of texture feature in the features vector, and \max is the maximum value of texture feature in the features vector. In this study, the extracted features will be normalized using this method to the range $[0, 1]$ (Tantisatirapong, 2015).

2.4.2.2 Z-Score Normalization

Z-Score normalization is widely used in statistical analysis, and is also known as zero mean normalization. The values of feature vector are normalized based on the mean and standard deviation of features vector as defined in Eq. 2.12 (Han et al., 2011; Larose, 2005; Takayanagi et al., 2011; Jayalakshmi and Santhakumaran, 2011).

$$v' = \frac{v - m}{\sigma} \quad 2.12$$

where m and σ are the mean and standard deviation respectively of the feature vector. This method is preferable when the actual minimum and maximum values of the feature vector are unknown (Han et al., 2011).

2.4.2.3 Decimal Scaling Normalization

Decimal scaling normalization is performed by moving the decimal point of values of texture features. The number of decimal points moved depends on the maximum absolute value of the features vector, as defined in Eq. 2.13 (Han et al., 2011).

$$v' = \frac{v}{10^j} \quad 2.13$$

where j is the smallest integer and $Max(v') < 1$.

2.4.2.4 Median Normalization

Median normalization uses the median value for each features vector to normalize each feature. It is a useful method when there is a need to compute the ratio between two hybridized samples, and the median value is not effected by the magnitude of extreme deviation. It is determined using Eq. 2.14 (Jayalakshmi and Santhakumaran, 2011).

$$v' = \frac{v_i}{median(v)} \quad 2.14$$

where v denotes the input features vector, and v_i is the required sample in the feature space.

2.4.2.5 Sigmoid Normalization

Sigmoid normalization is a non-linear method used to normalize the extracted features within range of 0 to 1 or -1 to +1, as defined in Eq. 2.15 (Jayalakshmi and Santhakumaran, 2011).

$$v' = \frac{e^v - e^{-v}}{e^v + e^{-v}} \quad 2.15$$

where v denotes the input features vector, and v_i is the required sample in feature space.

2.5 Classification

Classification is the process of sorting objects in images into separate classes and plays an important role in medical imaging, especially in the detection and recognition tumours and many other applications such as robotic and speech recognition (Dougherty, 2009). Classification is concerned with identifying criteria that can be used to discriminate different objects that may appear in images (Russ, 1990). Data classification includes two essential steps; in the first step, the classifier is built describing a predetermined set of data classes. This phase is known as the learning phase or training phase (Sonka et al., 2014; Han et al., 2011). Once the class labels of the given data has been assigned, we say that the classifier is supervised learning, otherwise if the class labels are unknown, the classifier is said to be unsupervised (Han et al., 2011; Larose, 2005; Solomon and Breckon, 2011). In the second step, the classifier performance is assessed on a new set of data know as test data and this phase is known as the “test phase”. The classification techniques are grouped into statistical approaches and intelligent approaches. The statistical approaches are based on computing the probability distributions and estimate some parameters such as the mean and standard deviation to provide a better representation of the classes such as LDA, SVM and k-nearest neighbour (KNN) (Li and Ogihara, 2006; Han et al., 2011; Dougherty, 2009; Dubitzky et al., 2007). Intelligent approaches include learning capabilities and involves to use an artificial intelligence techniques in the classification process, such as ANN (Hagan et al., 1996; Lekutai, 1997). A number of common classification methods have been used to distinguish and differentiate brain tumours in MRI images and the three most popular classifiers are LDA (Zacharaki et al., 2009; Takayanagi et al., 2011), SVM with radial basis function (Nagarajan et al., 2013; Kharrat et al., 2010; Bauer et al., 2011; Hackmack et al., 2012; Kalbkhani et al., 2013) and MLP (Pantelis, 2010; Saritha et al., 2013; Antkowiak, 2006).

2.5.1 Linear Discriminant Analysis

The linear discriminant analysis (LDA) is a classification method originally developed by Fisher in 1936, it is based on searching for a linear combination of the variables that best discriminate among classes (Li and Ogihara, 2006). The classification is implemented on the transformed space based on maximizing the ratio of between-class variance to the within-class variance to find the optimal vector Φ to separate the classes. It computes the probability distributions and derives the mean and standard deviation to provide a better representation of

classes. It is commonly used for data classification and dimension reduction (Li and Ogihara, 2006; Chen and Tian, 2011; Balakrishnama and Ganapathiraju, 1998). Let us assume that there is a set of m p -dimensional samples x_1, x_2, \dots, x_m (where, $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$) belonging to two different classes C_1 and C_2 . The with-in classes scatter matrix S_w is computed using Eq. 2.16 (Li and Ogihara, 2006; Zacharaki et al., 2009; Mihelič and Žibert, 2008).

$$S_{w_i} = \sum_{i \in c} (x_i - \mu_i)(x_i - \mu_i)^T \quad 2.16$$

where i is the number of classes and μ_i is the overall mean of points within the same class that is calculated using Eq. 2.17.

$$\mu_i = \frac{1}{N_i} \sum_{i \in c} x_i \quad 2.17$$

where N is the number of samples in C .

Similarly, the difference between the classes' means is known as the between-class scatter matrix S_b and it is computed using Eq. 2.18.

$$S_b = \sum_c (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T \quad 2.18$$

where μ_1 and μ_2 are the mean of all points in the first and second class respectively.

As defined earlier, the main objective of LDA is to maximize the distance between the classes by increasing the ratio of between-class scatter to the within-class scatter as shown in Fig. 2.4. This ratio is known as the Fisher criterion and is defined in terms of S_w and S_b as given in Eq. 2.19 (Zacharaki et al., 2009).

$$\max_{\Phi} J(\Phi) = \frac{\Phi^T S_b \Phi}{\Phi^T S_w \Phi} \quad 2.19$$

By taking the derivative of $J(\Phi)$, then it is easy to find a vector Φ that maximizes $J(\Phi)$.

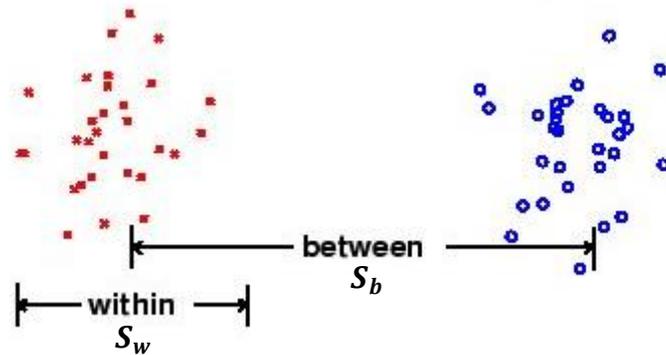


Figure 2.4: Separation between two classes.

Finally, each group of features will have a normal distribution of discriminant scores, the success of separation and classification depends on the degree of overlapping between the discriminant score distributions. A poor classification happens when the discriminant scores overlap widely, and a good classification happens when the discriminant scores do not overlap meaning that misclassification will be minimal as shown in Fig. 2.5.

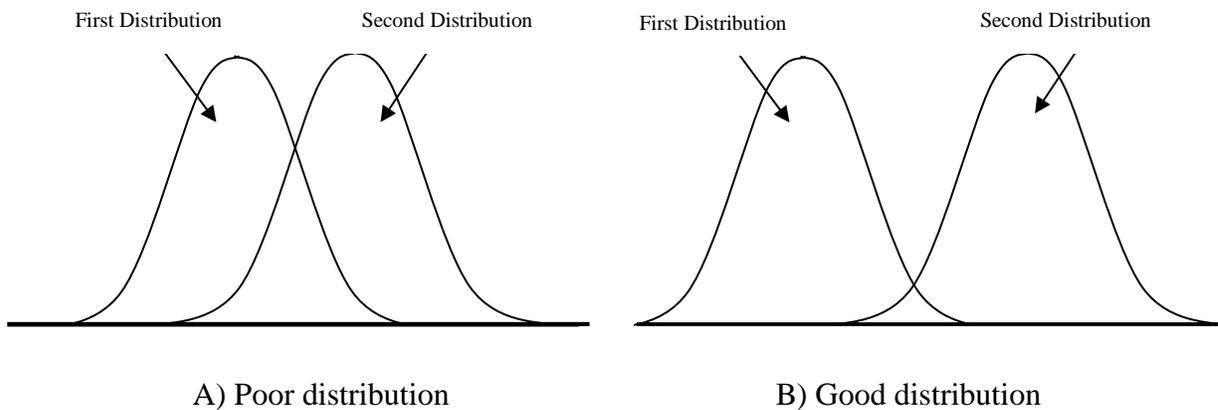


Figure 2.5: The discriminant distributions.

2.5.2 Support Vector Machines

The support vector machine (SVM) was developed in 1992 by Vapnik and colleagues Boser and Guyon. It is considered as the state-of-the-art supervised learning model and it has been used in various application such as handwriting recognition, object recognition, speaker identification and medical diagnosis (Han et al., 2011; Tantisatirapong, 2015). It is a statistical

approach for performing both linear and nonlinear classification by transforming the features into a higher dimension using a nonlinear mapping. It searches for the linear or nonlinear optimal separating hyperplane to separate the features into classes (Han et al., 2011). SVM provides a robust classification framework that works efficiently for situations with a moderately large number of features and relatively small sample size (Dubitzky et al., 2007). The optimal separating hyperplane is constructed when the maximum distance between the two closest data points in the two classes is found as shown in Fig. 2.6 (Han et al., 2011; Dubitzky et al., 2007). The decision function d is defined by Eq. 2.20.

$$d(x, \omega, b) = \omega \cdot x + b = \sum_{i=1}^n \omega_i x_i + b \quad 2.20$$

where x is the attributes vector, ω is the weight vector, b is a scalar and represents a bias or also known as an additional weight w_0 , and n is the number of attributes. If there are two attributes (A_1, A_2), then $X = (x_1, x_2)$, where x_1 and x_2 are the values' vectors of these two attributes. The separating hyperplane function can be estimated from the decision function when sets it equal to zero as given in Eq. 2.21.

$$\omega \cdot x + \omega_0 = \sum_{i=1}^n \omega_i x_i + \omega_0 = 0 \quad 2.21$$

Therefore, any point that lies above the separating hyperplane will satisfy Eq. 2.22:

$$\omega_0 + \omega_1 x_1 + \dots + \omega_n x_n > 0 \quad 2.22$$

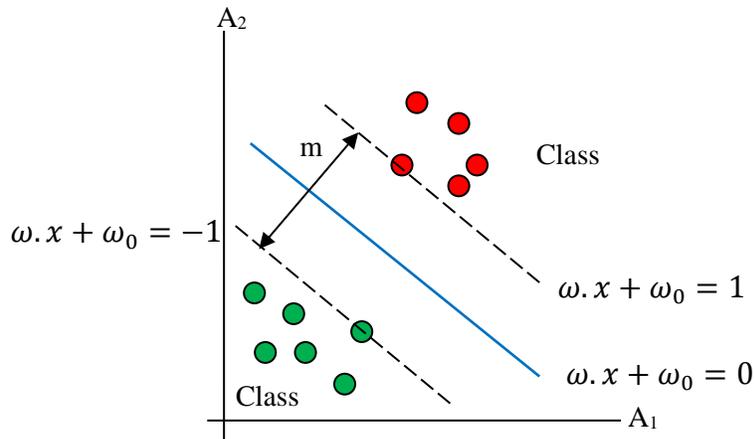


Figure 2.6: Two classes separated by a hyperplane.

Similarly, any point that lies below the separating hyperplane will satisfy Eq. 2.23:

$$\omega_0 + \omega_1 x_1 + \dots + \omega_n x_n < 0 \quad 2.23$$

While the expression of the margins “sides” can be written as Eq. 2.24 and Eq. 2.25:

$$\text{H1: } \omega_0 + \omega_1 x_1 + \dots + \omega_n x_n \geq 1 \quad 2.24$$

$$\text{H2: } \omega_0 + \omega_1 x_1 + \dots + \omega_n x_n \leq 1 \quad 2.25$$

where H1 and H2 are the hyperplanes also called support vectors.

By combining Eq. 2.24 and Eq. 2.25, Eq. 2.26 is obtained.

$$\omega_0 + \omega_1 x_1 + \dots + \omega_n x_n \geq 1, \forall_i \quad 2.26$$

Any attribute sample which falls on the support vectors H1 or H2 will satisfy Eq. 2.26 and it becomes as Eq. 2.27:

$$\omega_0 + \omega_1 x_1 + \dots + \omega_n x_n = 1 \quad 2.27$$

Consequently, it is easy to find the size of the maximal margin m , which represents the distance from any point that is located on the support vector H1 to the hyperplane is $\frac{1}{\|\omega\|}$, where the $\|\omega\|$ is the Euclidean norm of ω , which it is $\sqrt{\omega_1^2 + \omega_2^2 + \dots + \omega_n^2}$, and this is equal to the distance from any point which is located on the support vector H2 to the hyperplane. Thus, the maximal distance between the support vectors is $\frac{2}{\|\omega\|}$.

In order to get the maximum separability, the maximal distance should be maximized $\frac{2}{\|\omega\|}$ or minimized $\|\omega\|$. The latter term is equivalent to minimizing $\frac{1}{2} \|\omega\|^2$ and it could be solved by using the Lagrangian method by minimizing Eq. 2.28 (Sonka et al., 2014; Hamel, 2009).

$$L = \frac{1}{2} \|\omega\|^2 - \sum_n \alpha_n [y_n(\omega \cdot x_n + \omega_o) - 1] \quad 2.28$$

By deriving Eq. 2.28, Eq. 2.29 is obtained:

$$\frac{\partial L}{\partial \omega} = \omega - \sum_n \alpha_n y_n x_n = 0 \quad 2.29$$

and ω is given in Eq. 2.30

$$\omega = \sum_n \alpha_n y_n x_n \quad 2.30$$

By substituting ω in Eq. 2.28, Eq. 2.31 is obtained.

$$\frac{\partial L}{\partial \omega_o} = - \sum_n \alpha_n y_n = 0 \quad 2.31$$

Then the discriminant function of SVM classifier can be deduced as Eq. 2.32.

$$d(x^T) = \sum_n y_n \alpha_n x_n x^T + \omega_o \quad 2.32$$

where y_n is the class label of support vector x_n , x^T is the set of attributes and α_n is a numeric parameter which is determined automatically by the optimization process (Han et al., 2011; Hamel, 2009).

2.5.3 Artificial Neural Network

The artificial neural network (ANN) is a technique that is used to seek and build an intelligent model that simulates the working network of the neurons in the human brain (Ross, 2009). It has been successfully used to solve many problems. There are many types of ANNs that are designed to address a wide range of problems in the area of pattern recognition and classification, signal processing, object recognition and robotics. These types of ANNs are determined by the processing elements, pattern of connectivity, the strength of weights and training or learning rules which specify an initial set of weights and how these weights adapt during implementation (Lekutai, 1997; Birry, 2013; Wilson and Ritter, 2000).

2.5.3.1 Multilayer Perceptron

The multilayer perceptron (MLP) neural network model is well known and most popular among all ANNs. It has been widely used to provide a nonlinear mapping between its input and output to solve different kinds of problems such as prediction and diagnosing of many medical applications (Birry, 2013; Jiang et al., 2010; Hu and Hwang, 2001). Typically, the MLP network configuration as shown in Fig. 2.7, has an input layer, one or more hidden layers and an output layer. Such that the size of input layer corresponds to the number of descriptors in features vector. Usually, there is no neuron function implemented in that layer (Hu and Hwang, 2001). While the number of neurons in output layer depends on the particular problem at hand. The neurons of MLP are fully connected, that's mean that every neuron in a given layer is connected to all neurons in next layer (Larose, 2005; Han et al., 2011; Günther and Fritsch, 2010). Typically, there is no theoretical limit for choosing the number of hidden layers, but practically most purposes of pattern recognition algorithms use one or two hidden layers and maximally three hidden layers are acceptable to solve problems of any complexity. Increasing the number of hidden layer does not benefit to improve the accuracy of the classification process. It may lead to overfitting, memorizing the training set and increasing computation time. While increasing the number of neurons in hidden layer leads to increase the power and flexibility of the network for tackling complex problems. Therefore, the number of hidden layers and number of neurons in each hidden layer are chosen experimentally by user. If overfitting is occurred, reducing the number of neurons in the hidden layer may considered is an important (Larose, 2005; Birry, 2013). It has been proven that it is possible to approximate an arbitrarily complex mapping within a finite support by using two hidden layers and a sufficient number of neurons in hidden layer (Hu and Hwang, 2001).

Many activation functions have been used with neural network but only a few have been used in practical application. The most common activation functions that are used widely in decision-making neurons for classification and pattern recognition tasks are the sigmoid function and hyperbolic tangent function. Such that, they are satisfying the approximation conditions of ANNs, nonlinear functions and more differentiable (Lekutai, 1997; Özkan and Erbek, 2003; Negnevitsky, 2005). The differentiable activation function makes the function that is computed by a neural network is differentiable (Rojas, 1996).

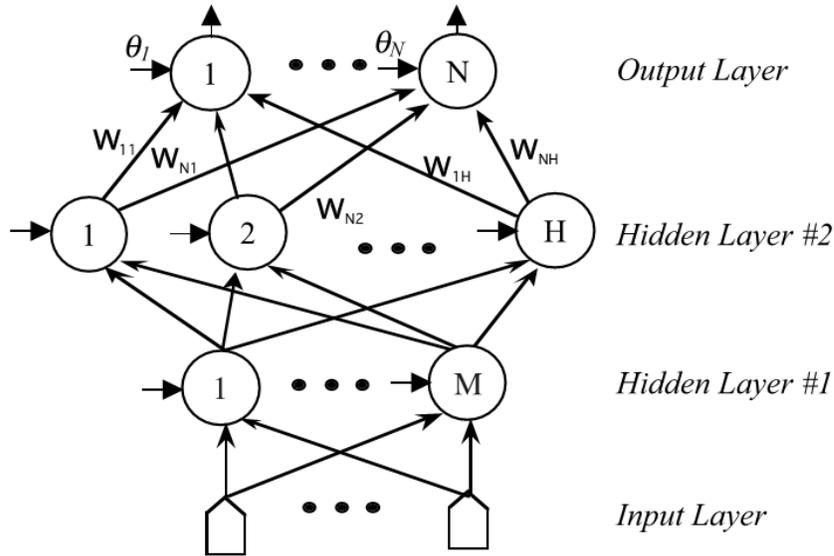


Figure 2.7: Multilayer perceptron configuration (Hu and Hwang, 2001).

The sigmoid function is computed as Eq. 2.33.

$$f(x) = \frac{1}{1 + e^{-x}} \quad 2.33$$

This function has ability to map a large input domain of any value between plus and minus infinity, into the small range of 0 to 1. Usually neurons with this function are used in the back-propagation networks (Negnevitsky, 2005). Because of it is nonlinear function, it allows the MLP to classify the data that it is linearly inseparable (Han et al., 2011). As well as, the error function that is produced by a sigmoid function is smooth or flat and it is always has a positive derivative (Birry, 2013; Rojas, 1996).

The second most widely used function is the hyperbolic tangent function, it is computed as following Eq. 2.34 (Graupe, 2013):

$$f(x) = \tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad 2.34$$

It is a bipolar version of the sigmoid function and used to map data onto the range of -1 to +1. Such that, the -1 and +1 output values represent minus and plus infinity respectively. The hyperbolic tangent function has faster convergence of learning algorithm than sigmoid function in addition it is more efficient for the classification (Özkan and Erbek, 2003).

The most popular and successful learning algorithm for training the MLP is the back propagation algorithm. It was developed by Rumelhart Hinton and Williams in 1986 (Lekutai, 1997). It is based on using an iterative descent method (Larose, 2005) for minimizing mean squared error between target and the actual output of MLP network. There are several issues associated with designing and training a MLP network such as number of hidden layers and number of neurons in hidden layers. These issues will be discussed in section 2.5.3.2.

2.5.3.2 Determining Number of Neurons in the Hidden Layers of MLP

Deciding the number of neurons in the hidden layers is an important issue because the hidden layers do not deal immediately with the external environment. They have a vital influence on the final output of the network. Therefore, both the number of hidden layers and the number of neurons in the hidden layers must be chosen carefully (Panchal et al., 2011). There are some guidelines are followed for determining the number of neurons in the hidden layers (Xu and Chen, 2008; Gunasekara et al., 2009):

- The number of neurons in the hidden layer should be in between the size of the input layer and the size of the output layer.
- The number of neurons in the hidden layer should be around $2/3$ of the input layer size, plus size of the output layer.
- The number of neurons in the hidden layer should be less than twice the size of the input layer.

2.6 Image Segmentation

The principle goal of the segmentation process is partitioning an image into meaning and homogeneous regions with respect to one or more characteristics (Solomon and Breckon, 2011; Dougherty, 2009; Sonka et al., 2014; William, 2001; Russ, 1990). Subdivision levels of segmentation depend on the problem being solved (Gonzalez and Woods, 2002). In medical imaging, segmentation is an important tool for feature extraction, image measurements and image display. It has been useful in many applications such as coronary border in angiograms, multiple sclerosis lesion quantification, surgery simulations, surgical planning and measuring tumour volume. Segmentation is useful to classify image pixels to different anatomical regions, such as bones, muscles and blood vessels. Furthermore, it is used to classify the pixels of pathological regions, such as cancer, tissue deformities and multi sclerosis lesions (Ho et al., 2002; Dvořák et al., 2013; Prastawa et al., 2009; Bankman, 2009). The segmented image should have the following aspects in order to get a good image segmentation;

1. The regions of the segmented image should be uniform.
2. The internal area of the regions should be clear without small holes.
3. The adjacent regions of the segmented image should have a significant difference.
4. The boundary of the segmented regions should be smooth and not coarse (William, 2001).

Segmentation techniques are divided into manual, semi-automated and fully automatic based on the degree of user interaction. The manual segmentation depends on the domain-knowledge of the user and the resulting output from an expert. This is known as the ground truth or the gold standard. Fully automatic segmentation does not need for user interaction and it requires less processing time but it is likely to perform less satisfactorily on medical images due to the complexity and inhomogeneity of anatomical texture. Semi-automatic segmentation method is used when the pathological area can be easier to identify visually but not automatically, then the user's supervision is integrated with computer algorithm to achieve an optimal segmentation of the region of interest (RoI) (Tantisatirapong, 2015). A wide variety of brain tumour segmentation techniques have been proposed and still now there is no standard segmentation technique that can produce satisfactory results (Gordillo et al., 2013). There are many segmentation techniques of very different nature each one has benefits and drawbacks.

Different applications need different segmentation techniques and the main factors that could be used to choose an appropriate segmentation techniques are (Rousseau, 2009):

- 1- Level of noise in the image.
- 2- Characteristics of the required object such as topology, homogeneity and size.
- 3- Sharpness of the object's contour.
- 4- Image contrast.

In the last few decades the number of publications that are devoted to automated brain tumour segmentation has grown rapidly due to the progress in the medical imaging field (Menze et al., 2015). Prastawa et al. (2004) described a framework for automated brain tumour segmentation of MRI brain scans where the detection of edema was implemented simultaneously with tumour segmentation. This framework is composed of three stages of analysis of T2-w images; detection of brain abnormalities by using atlas registration, determination of the presence of edema within the abnormality (tumour), and generation of geometric and spatial constraints to discriminate tumour and edema regions. Parametric deformable models, also known as active contour models or snakes, are shown to be strongly suitable for determining these boundaries (Gordillo et al., 2013). These approaches enable segmentation, matching and tracking of anatomical areas by exploiting conditions derived from anatomical and biological knowledge about location, size and shape of the anatomical areas (Nabizadeh, 2015). These deformable models are defined as curves or surfaces that move under the influence of weighted internal and external forces. Internal forces are responsible for the smoothness of the curves, while external forces are responsible for pushing and pulling the curves toward the boundaries of anatomical area. An automated brain tumour and edema segmentation algorithm was proposed to implement a fast segmentation of MRI brain slices based on bounding boxes method (Saha et al., 2012). The Bhattacharya coefficient of grey scale intensity histograms has been used as a score function to measure the similarity between two intensity histograms to locate bounding boxes around the abnormal area in MRI brain slices. This method parallelizes searching for the most dissimilar region in the MRI brain slice between the left and right hemispheres of the brain in the axial view of the MRI scan (Saha et al., 2012; Ray et al., 2008b). An automated algorithm for detecting the tumour location in a single MRI brain slice and identifying the tumour boundaries by using an unsupervised learning algorithm called Force algorithm, was proposed

in (Khandani et al., 2009). A set of prior operations for skull and non-tumour pixel removal, using histogram analysis and exponential transformation, was applied with subsequent segmentation of the tumour area by histogram thresholding. Mikulka and Gescheidtov (2013) proposed an automated segmentation algorithm to recognize brain tumours, edema and necrosis in T1-w and T2-w images. A binary mask was created to enable measurement in perfusion weighted images, where the tumour type could be recognized by detecting the level of perfusion of contrast agent in the pathological tissue. Havaei et al. (2017) presented a fully automatic brain tumour segmentation method based on convolutional neural network (CNN), that exploits both local features as well as more global contextual features simultaneously. In this approach segmentation was performed slice by slice from the axial viewing and each pixel was associated with different MR modalities; T1-w, T2-w, T1c-w and FLAIR images. The CNN architecture included several layers that are stacked on top of each other and were used to form a hierarchy of features called a feature map, corresponding to neurons within a neural network. The reported maximum achievable accuracy of this approach was 88%, and the specificity and sensitivity were 89% and 87% respectively. Previous studies commonly focus on segmenting each slice individually (“slice-by-slice”), then merge these to obtain a three-dimensional volume or a continuous surface. However, the resulting segmentation can include inconsistencies and non-continuous surface due to missing important anatomical information in three-dimensional space. These approaches do not exploit all the features of the full MRI slices (Mikulka and Gescheidtov, 2013; Despotović et al., 2015). Ho et al. (2002) developed a new method for the automatic segmentation of anatomical structures from volumetric T2-w and T1c-w MR brain scans by using a three-dimensional level-set and a probability map. The probability map was used locally to guide the propagation direction and speed of a level-set snake as well as to derive an automatic initialization of the snake. The snake can grow inside the tumour area even if the initialization seed point covers a small portion of the tumour. Rousseau (2009) proposed an automated algorithm for heart segmentation, known to be a difficult task due to the complexity of the organ shape and its topology. A three-dimensional deformable model approach was used as a best segmentation method that does not require training data. As well as, it requires a careful initialization with an initial contour that is close to the object of interest. The approach also requires user guidance to place landmarks in the image to steer the segmentation. Klotz (2013) proposed an automated algorithm to determine the thickness of the retinal nerve fibre layer

which can serve as an early indicator of glaucoma. The two-dimensional active contour without edge (2DACWE) and three-dimensional active contour without edge (3DACWE) were used to segment synthetic and real human retina. All the above mentioned active contour segmentation approaches require the estimation of the initial close snake contour for the tumour (Loizou et al., 2009). Alternatively, an automated three-dimensional segmentation algorithm based on kernel-based fuzzy C-means (KFCM) was proposed to separate the MRI breast slices into different parts by Song et al. (2013). However, no evaluation in terms of accuracy of their method was presented. Another use of fuzzy rules for MRI brain scans was reported in (Dou et al., 2007).

In general, to distinguish segmentation techniques, they are classified into groups based on the image information which is used to implement the segmentation. These are described in the following subsections.

2.6.1 Pixel Based Segmentation

This type of segmentation is also known as threshold-based methods. They are conceptually the simplest approaches of segmentation and commonly used in two dimensional images. They consider only intensity value of the current pixel and discarding its neighbouring pixels. Actually, it is noted that these methods are not segmentation procedures because each pixel is isolated from its neighbourhood. Therefore, it cannot be guaranteed that actually only the connected segments are obtained. Most of pixel based techniques depend essentially on measuring thresholds from the histogram of an image (Petrou, 2011; Dougherty, 2009; Naji et al., 2012; Watt and Policarpo, 1998). In these methods, the objects of the image are classified by comparing their pixels' intensities with one or more intensity thresholds. These thresholds can be either global or local. If the object can be separated from the background of the image by a single threshold this is named as global thresholding. However, if there are more than two objects, then the segmentation should be implemented using local thresholding (Jin et al., 2014; Gordillo et al., 2013; Naji et al., 2012; Russ, 1990). Automated threshold selection is essentially based on considering the image histogram, such that the histogram typically includes two peaks; one corresponding to the pixels of the objects and the other one to the pixels of the background. The threshold needs to be chosen so that these two peaks are clearly separated from each other (Solomon and Breckon, 2011). The major problem with thresholding is that only the intensity

information is considered and the relationships between the pixels are ignored. There is no guarantee that the segmented pixels by thresholding are contiguous. As well as it probably includes extraneous pixels that do not attend to the desired region or attend to the background region. Sometimes isolated pixels within the region especially near the boundaries of the region are ignored. Generally, threshold-based segmentation methods are unable to exploit all the information that is provided by MRI slice and in most cases are usually used to separate and eliminate the background of MRI slice (Gordillo et al., 2013; Wilson and Ritter, 2000; Russ, 1990).

2.6.2 Region Based Segmentation

This type segmentation is considered as one of the conceptually simplest approaches for image segmentation. It is based on dividing the image into regions according to predefined similarity criteria. It is also called region merging and starts with a single pixel or a group of pixels that are called seeds. Neighbours of the seed are checked and those satisfying the similarity criteria are added to the same structure of interest (Solomon and Breckon, 2011). The similarity between pixels can be based on intensity information and/or edges in the image (Pham et al., 1998). The procedure repeats until no more pixels can be added to the structure of interest. The similarity criteria are determined according to the features of the image and the seeds can be chosen manually by the user or automatically by a computer algorithm. The main characteristic of region growing is that it is capable to segment regions that have similar properties and generating connected region (Rogowska, 2009). The main disadvantage of region growing methods is the PV effect which limits the accuracy of MR brain image segmentation. Such that, PV blurs the borders between different tissues because the voxel may contain more than one kind of tissue types (Gordillo et al., 2013; Sato et al., 2000; Jin et al., 2014). It is more sensitive to noise, causing extracted regions to have holes or become disconnected (Pham et al., 1998). Additionally, if the seed point is not properly chosen, the region grows outside the object of interest or merges with another region that does not belong to the desired object (Dougherty, 2009).

2.6.3 Edge Based Segmentation

It is based on finding the differences instead of similarity between pixels to determine the close boundaries which correspond to objects of an image (Dougherty, 2009). It is computationally fast and does not need any prior information about image content. Usually, it operates on edge magnitude and/or phases' images that are produced by an edge operator suited to the expected characteristics of the image (Rogowska, 2009). It has been developed to be strongly sensitive to the significant variations in grey level values and decide whether or not a pixel lies on an edge independently of the neighbouring pixels (Birry, 2013). The variations in a grey level value mean that is gradually changing from the background to the object value. This approach can be used to avoid a bias in the size of the segmented object without a complex thresholding scheme because it is based on the fact that the position of an edge is given by an extreme of the first order derivative of the image function (Jähne, 2005). The main problems of edge based segmentation is that often the edges do not enclose the object completely. To solve this problem, an extra post-processing steps of linking or grouping edges that correspond to a single boundary is required to combine edges into edge chains that correspond better with edges in the image. They are more sensitive to image noise, unsuitable information in an image, and if the image's region features differ by only a small amount between regions, a detected edge may be broken (Sonka et al., 2014; William, 2001; Rogowska, 2009).

2.6.4 Deformable Model

The basic idea of the deformable model for image segmentation is to embed an initial contour into the image, then let it evolves according to constraints derived from the image together with a prior knowledge about the location of RoI. The contour continues to evolve under the influence of internal and external forces until stopping on the edge of RoI. It is also variously named snakes, active contour, balloon and deformable contour (Dougherty, 2009; Jähne, 2005). Among all segmentation techniques, the deformable model has proved to be a successful and efficient segmentation technique for a wide range of application especially medical application due to its capability of accommodating the often significant variability of biological structures over time and cross different patients (Rousseau, 2009; Tantisatirapong, 2015). Jin et al. (2014) and Gordillo et al. (2013) concluded that the good results of brain tumour segmentation using conventional methods (e.g. pixel based method, region based method and edge based method)

are hard to achieve. In most situations, these methods were used as a pre-processing step in the segmentation of brain tumour. Additionally, due to the appearance of volumetric three-dimensional medical imaging data, the segmentation of this data has become a challenging problem to extract boundary elements that belong to the same structure. Therefore, the deformable model was proposed to improve this problem by combining constraints which are derived from the image and a prior knowledge of the object such as location, shape and orientation. The deformable models were originally developed to address the problems in computer vision and medical image analysis. Two-dimensional and three-dimensional deformable models have been applied to segment, visualize, track and quantify a variety of anatomic structures such as the brain tumours, heart, face, cerebral, coronary and retinal arteries, kidney, lungs, etc. (McInerney and Terzopoulos, 2009). The deformable model is susceptible to fall in local minimum due to some the local features of the image, or when some of the object's edges are missing. Therefore the brain tumours segmentation with deformable model require a careful initialization with the initial contour that is close to the tumour's boundaries (Rousseau, 2009). Although there are several general segmentation methods such as thresholding (Petrou, 2011; Dougherty, 2009; Naji et al., 2012; Jin et al., 2014) and region growing (Fabijańska, 2009; Jin et al., 2014), they are not applicable on the brain lesions segmentation. Active shape model and active appearance model are avoided because these methods require a training set that consists of many images that are manually segmented. Similarly, a model based method is avoided because it requires atlas registration. Such that atlas is built from a given training set, for the same reason that the segmentation based on training set is not preferable and the registration remains challenging and yet not solved for the general use (Nabizadeh, 2015; Nabizadeh and Kubat, 2015). Furthermore the atlas registration introduces a bias in the segmentation because the algorithm searches for a shape similarity to one of the atlas (Rousseau, 2009). Therefore, the concentration in this study will be on using a deformable model approach (e.g. 2DACWE) due to its simplicity and robustness, appearance of three-dimensional MRI data, ability to extract boundary elements belonging to the same structure and integrate these elements into a single object. In addition to the ability of segmenting anatomic structures by using derived constraints from the image and a prior knowledge about the location of these structures and capability of segmenting the biological structures which are often significant variability over time and across different individuals.

Furthermore, the deformable model provides highly intuitive interaction mechanisms that allow clinicians and radiologists to bring their expertise to image interpretation task (Jin et al., 2014; Rousseau, 2009).

2.6.4.1 Active Contour

Active contour (AC) is the most popular method and was introduced by Kass et al. (1988). It is defined by an energy function and a solution is found using techniques of variation calculus and the finite difference methods where a set of curves are defined and deformed within digital images to recover object shapes (Loizou et al., 2009). It has been widely applied to image segmentation, analysis of dynamic image data or three-dimensional image data and medical images. There are several distinct advantages of AC models over other conventional segmentation methods. First, AC models can achieve sub-pixel accuracy of object boundaries. As well as it can be easily formulated under a principled energy minimization framework, and allow the incorporation of various prior knowledge such as shape and intensity distribution for robust image segmentation (Chunming et al., 2008). It generates snake or contour within an image domain, the contour can be moved and directed under the effect of its internal and external forces from the image data. The internal and external forces of the contour correspond to the object boundary and desired features in an image (Xi-ping et al., 2002; Pham et al., 1998). The contour does not solve the entire problem of finding objects in images. It depends on the user interaction and information from image data that is related with time or space. The interaction means that the approximate shape or the starting point of the contour somewhere inside the desired object should be specified in addition to image-based information. All these information are used to push the contour toward the desired object's contour (Sonka et al., 2014). The location of contour in the given image associates with the energy function. Mathematically, a deformable contour moves through an iterative process and adapts itself by a dynamic process that minimizes an energy function. Where, the energy function achieves the minimum when the contour reaches the object boundary within the image. Initially, the initial contour C is defined close to the edge of the desired object and it is guided by defining the corresponding energy function that is given by Eq. 2.35 (Rousseau, 2009; Kass et al., 1988; Xi-ping et al., 2002; Talebi et al., 2010; Xu and Prince, 1998; Sonka et al., 2014).

$$E(C) = E(\phi(x, y)) = w_1 E_{int}(\phi(x, y)) + w_2 E_{image}(\phi(x, y)) + w_3 E_{ext}(\phi(x, y)) \quad 2.35$$

where E is the energy function, w_1 , w_2 and w_3 are real positive values, and $\phi(x, y)$ is the level set that uses the signed distance function (SDF) and represents the mathematical description of C , and makes it always continuous and closed. E_{int} denotes the energy inside C that is used to control the rate of stretching and smoothing and preventing the discontinuity in C , as given in Eq. 2.36 (Sonka et al., 2014; Loizou et al., 2009).

$$E_{int}(\phi(x, y)) = \alpha |\phi(x, y)'|^2 + \beta |\phi(x, y)''|^2 \quad 2.36$$

where the first order term is the elasticity and is controlled by α , and the second order term is the stiffness and controlled by β . They are responsible for controlling the natural behaviour of C and makes it like a membrane. They can be adjusted by balancing this term against an external energy E_{image} from the image (Toennies, 2012).

Whereas, E_{image} is related to the image gradient and includes a combination of some relevant features such as the gradient of the edge, lines, regions and textures that attract the contour to follow high gradients in the image, as given in Eq. 2.37 (Loizou et al., 2009; Toennies, 2012).

$$E_{image}(\phi(x, y)) = \int_0^1 P(\phi(x, y)) dx dy \quad 2.37$$

where P is a scalar potential function defined on the image plane, and it is defined by given Eq. 2.38.

$$P(\phi(x, y)) = -c |\nabla [G_\sigma(x, y) * I(x, y)]|^2 \quad 2.38$$

where c controls the magnitude of potential function, $G_\sigma(x, y)$ is a two-dimensional Gaussian function whose characteristic standard deviation is σ , ∇ is the gradient operator and $*$ denotes the convolution process between the Gaussian function and given image $I(x, y)$.

While, E_{ext} is defined by the user and it is optional. Occasionally it includes two different functions that attract the snake to lines and edges. It is given in Eq. 2.39 (Sonka et al., 2014; Loizou et al., 2009).

$$E_{ext} = w_{line} E_{line} + w_{edge} E_{edge} \quad 2.39$$

The line function is given in Eq. 2.40 (Sonka et al., 2014).

$$E_{line} = I(x, y) \quad 2.40$$

where $I(x, y)$ denotes the image intensity at the specific locations (x, y) , and the sign of w_{line} specifies whether attraction of the snake is to the light or dark lines.

The edge function is given in Eq. 2.41 (Sonka et al., 2014).

$$E_{edge} = -|\nabla I(x, y)|^2 \quad 2.41$$

where ∇ is the gradient operator that attracts the snake to the large image gradients that have strong edges.

The main drawbacks of AC are sensitive to the initial conditions and the difficulties associated with the topological changes for merging and splitting of the evolved contour. Such that the initial contour should be closed to the desired boundary or it will likely converge to the wrong boundary (Thapaliya et al., 2013; Xu and Prince, 1998). Additionally, due to AC relying on image gradient to guide and stop contour evolution, the intensity homogeneity is essential to achieve efficient segmentation. In fact, intensity inhomogeneity occurs significantly in real medical images from different modalities due to technical limitations, artefacts introduced by the scanned object and variations in object susceptibility (Chunming et al., 2008). Additionally, reliance on the gradient method is bounded because it is sensitive to noise and some objects have boundaries that are not well defined through the gradient. This makes the energy function never approaching zero on the edges (Chan and Vese, 2001). Finally, there are series of slices in medical images where objects boundaries are not clear and region statistics change dramatically in some parts of the objects' boundaries between neighbourhood slices (Xi-ping et

al., 2002). All of these issues are addressed by the 2DACWE method (Chan and Vese, 2001), that it will be described in details in 2.6.4.3.

2.6.4.2 Level Set Method

The level set method was introduced in the late 80's by Sethian and Osher (Rousseau, 2009). It is a powerful tool for implementing contour evolution and managing automatic topology adaptation using only geometric measures and some prior information from image data to recover RoI boundaries (Sonka et al., 2014). It also provides the basis for a numerical scheme that is used by geometric deformable models. In level set method, an evolving contour C is implicitly represented as a zero level set of a higher-dimensional scalar function referring to the level set function $\phi(C)$ which is defined on the same domain of the image. Instead of tracking a contour C through time, the level set method evolves the contour by updating the level set function $\phi(x, y, t)$ at fixed coordinates through time. It remains a valid function while the embedded contour C can change its topology (Xu et al., 2000). Figure 2.8 shows how the contour C is embedded at zero level set in the evolving surface $\phi(x, y, t)$, and can be written as Eq. 2.42 (Rousseau, 2009; Chan and Vese, 2001; Crandall, 2009; Thapaliya et al., 2013; Xu et al., 2000; Parisot, 2013).

$$\phi(C, t) = 0 \tag{2.42}$$

By differentiating, Eq. 2.42 with respect to t and using the chain rule method, Eq. 2.43 is obtained (Xu et al., 2000; Rousseau, 2009).

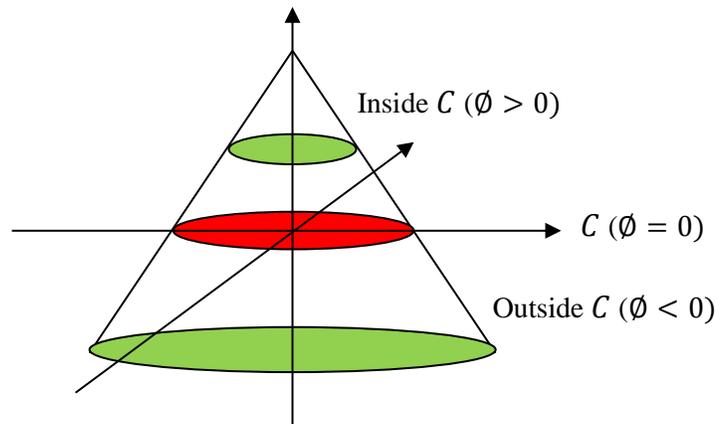


Figure 2.8: Evolving of contour C .

$$\frac{\partial \phi}{\partial t} + \nabla \phi \cdot \frac{\partial C}{\partial t} = 0 \quad 2.43$$

Accordingly, the unit normal to the level set curve is given by Eq. 2.44.

$$N = \frac{\nabla \phi}{|\nabla \phi|} \quad 2.44$$

Then, by using Eq.2.44, we can rewrite Eq. 2.43, as Eq. 2.45.

$$\frac{\partial \phi}{\partial t} + |\nabla \phi| N \cdot \frac{\partial C}{\partial t} = 0 \quad 2.45$$

Since the normal speed $V(k)$ is $N \cdot \frac{\partial C}{\partial t}$, then the evolution equation can be written as Eq. 2.46 (Parisot, 2013).

$$\frac{\partial \phi}{\partial t} + |\nabla \phi| V(k) = 0 \quad 2.46$$

Finally, the curvature k at zero level set measures how sharply the contour C bends or how quickly it changes direction, is given in Eq. 2.47 (Chan et al., 2000; Chan and Vese, 2001; Klotz, 2013).

$$k = \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) = \frac{\phi_{xx} \phi_y^2 - 2\phi_x \phi_y \phi_{xy} + \phi_{yy} \phi_x^2}{(\phi_x^2 + \phi_y^2)^{3/2}} \quad 2.47$$

where ϕ_x , ϕ_y , ϕ_{xx} , ϕ_{yy} and ϕ_{xy} are derivatives of level set function ϕ , they are derived by using a finite difference implicit scheme as given in Eq. 2.48, Eq. 2.49, Eq. 2.50, Eq. 2.51 and Eq. 2.52.

$$\phi_x = \frac{\phi_{x+\Delta x, y} - \phi_{x-\Delta x, y}}{2\Delta x} \quad 2.48$$

$$\phi_y = \frac{\phi_{x, y+\Delta y} - \phi_{x, y-\Delta y}}{2\Delta y} \quad 2.49$$

$$\phi_{xx} = \frac{\phi_{x+\Delta x,y} - 2\phi_{x,y} + \phi_{x-\Delta x,y}}{\Delta x^2} \quad 2.50$$

$$\phi_{yy} = \frac{\phi_{x,y+\Delta y} - 2\phi_{x,y} + \phi_{x,y-\Delta y}}{\Delta y^2} \quad 2.51$$

$$\phi_{xy} = \frac{\phi_{x+\Delta x,y+\Delta y} - \phi_{x-\Delta x,y+\Delta y} - \phi_{x+\Delta x,y-\Delta y} + \phi_{x-\Delta x,y-\Delta y}}{4\Delta x\Delta y} \quad 2.52$$

The segmentation by using two-dimensional contours has many problems especially when dealing with cross-sectional images such as MRI scan. Where, there is information loss because of neglecting the third dimension, broken boundary between slices and overlapping intensities (Aloui and Naceur, 2011). In order to avoid these issues, a three-dimensional level set method is used for segmenting a volumetric data where the curvature k at zero level set is given in Eq. 2.53 (Klotz, 2013; Sonka et al., 2014; Hasan et al., 2016b).

$$k = \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) = \frac{\sqrt{(\phi_{zz}\phi_y - \phi_{yy}\phi_z)^2 + (\phi_{xx}\phi_z - \phi_{zz}\phi_x)^2 + (\phi_{yy}\phi_x - \phi_{xx}\phi_y)^2}}{(\phi_x^2 + \phi_y^2 + \phi_z^2)^{3/2}} \quad 2.53$$

where ϕ_x , ϕ_y , ϕ_z , ϕ_{xx} , ϕ_{yy} and ϕ_{zz} are derivatives of the level set function ϕ , they are derived by using a finite difference implicit scheme as given in Eq. 2.54, Eq. 2.55, Eq. 2.65, Eq. 2.57, Eq. 2.58 and Eq. 2.59:

$$\phi_x = \frac{\phi_{x+\Delta x,y,z} - \phi_{x-\Delta x,y,z}}{2\Delta x} \quad 2.54$$

$$\phi_y = \frac{\phi_{x,y+\Delta y,z} - \phi_{x,y-\Delta y,z}}{2\Delta y} \quad 2.55$$

$$\phi_z = \frac{\phi_{x,y,z+\Delta z} - \phi_{x,y,z-\Delta z}}{2\Delta z} \quad 2.56$$

$$\phi_{xx} = \frac{\phi_{x+\Delta x,y,z} - 2\phi_{x,y,z} + \phi_{x-\Delta x,y,z}}{\Delta x^2} \quad 2.57$$

$$\phi_{yy} = \frac{\phi_{x,y+\Delta y,z} - 2\phi_{x,y,z} + \phi_{x,y-\Delta y,z}}{\Delta y^2} \quad 2.58$$

$$\phi_{zz} = \frac{\phi_{x,y,z+\Delta z} - 2\phi_{x,y,z} + \phi_{x,y,z-\Delta z}}{\Delta z^2} \quad 2.59$$

To implement the segmentation, an initial level set function (Eq. 2.42), must be defined and it is frequently based on SDF from each point in the grid to the zero level set at level set function with a sign depending on being inside or outside C , as given in Eq. 2.60 (Sonka et al., 2014):

$$\phi(C, t = 0) = \text{SDF} \quad 2.60$$

where,

$$\text{SDF} = \begin{cases} > 0 & \text{inside } C \\ = 0 & \text{on } C \\ \leq 0 & \text{outside } C \end{cases}$$

Since the stopping criteria is based on the speed function $V(k)$ that uses the image gradient to stop contour evolution. Considering the region properties of the segmented objects is essential and helpful. Therefore, we will focus on this study to use a piecewise constant minimal variance criterion based on the Mumford-Shah functional (Sonka et al., 2014) proposed by Chan and Vese (2001) to deal with such situations.

2.6.4.3 2D Active Contour without Edge

The two-dimensional active contour without edge (2DACWE) method, also known as Chan-Vese model, is an example of a geometric active contour model (Chan and Vese, 2001). The initial contour is evolved using a level set method and does not rely on the gradient of the image for stopping process. It integrates the statistical information which is related to the wanted regions to improve the quality of segmentation and is based on minimizing the Mumford-Shah function. The Mumford-Shah function is a function that is used to establish an optimum criterion for segmenting an image into sub-regions and suggested the energy function given in Eq. 2.61. It can be used for segmenting an image I into non-overlapping regions (Rousseau, 2009; Chan and Vese, 2001; Getreuer, 2012; Chunming et al., 2008; Pock et al., 2009).

$$\mathcal{F}^{MS}(u, C) = \lambda \int_{\Omega} (u - I)^2 dx dy + \int_{\Omega/C} |\nabla u|^2 dx dy + \mu \text{length}(C) \quad 2.61$$

where u denotes a set of piecewise smoothing functions that approximate the original image I , and smooths each of the connected components in the image domain Ω separated by the contour C . Therefore, 2DACWE can detect object boundaries both with and without gradients. For instance objects that are very smooth, or even have discontinuous boundaries (Chan and Vese, 2001; Crandall, 2009). To overcome the time complexity of solving the general Mumford-Shah function, it is required to suppose u to be constant on each connected component (Getreuer, 2012). An active contour approach was proposed by Chan and Vese (2001) based on minimizing Mumford-Shah functional by penalizing the enclosed area assuming that u is supposed to have only the two values which are given in Eq. 2.62 (Getreuer, 2012; Chunming et al., 2008).

$$u(x, y) = \begin{cases} c_1 & \text{where } x, y \text{ are inside } C \\ c_2 & \text{where } x, y \text{ are outside } C \end{cases} \quad 2.62$$

where c_1 and c_2 are the values of u inside and outside C respectively. The Chan-Vese energy function is given in Eq. 2.63 (Getreuer, 2012; Chunming et al., 2008; Chan and Vese, 2001; Thapaliya et al., 2013).

$$\begin{aligned} \mathcal{F}^{CV}(C, c_1, c_2) = & \mu \text{Length}(C) + v \text{Area}(\text{inside}(C)) + \lambda_1 \int_{\text{inside}(C)} |I(x, y) - c_1|^2 dx dy \\ & + \lambda_2 \int_{\text{outside}(C)} |I(x, y) - c_2|^2 dx dy \end{aligned} \quad 2.63$$

The regularity is controlled by penalizing the length in the first term, and the size is controlled by penalizing the enclosed area of C in the second term. These terms are called regularizing terms and are given in Eq. 2.64 and Eq. 2.65, and encourage the contour C to be smooth and short, and can be written by using level set form ϕ (Chan and Vese, 2001; Klotz, 2013) as.

$$length(\phi(x, y)) = \int_{\Omega} \delta_o(\phi(x, y)) |\nabla \phi(x, y)| dx dy \quad 2.64$$

$$area(\phi(x, y)) = \int_{\Omega} H(\phi(x, y)) dx dy \quad 2.65$$

$\lambda_1, \lambda_2, \mu \geq 0$ and $\nu \geq 0$ are fixed parameters controlling selectivity, where the energy function is minimized by fixing these parameters optimally. Meanwhile, λ_1, λ_2 control the internal and external forces respectively. These terms usually hold the same constant and hence a fair competition between these two forces (Chunming et al., 2008). Generally, $\lambda_1 = \lambda_2 = 1$ (Chan and Vese, 2001; Nixon and Aguado, 2008). Meanwhile, μ controls the smoothness of contour C and assumes a scaling role. However, the parameter is not constant across all experiments. If μ is large, only larger objects with smooth boundaries are segmented. If μ is small, objects of smaller size are segmented accurately (Tai et al., 2005; Chan and Vese, 2001; Getreuer, 2012). Typically, μ depends on image resolution (IRE), where $\mu = 0.1 \times IRe^2$ (Nixon and Aguado, 2008). Rousseau (2009) concluded that there is no automatic way to set these parameters and they are usually tuned by letting them vary and observing the results. Once the object has sharp boundaries, only a small influence of the curvature term is needed, meaning that the ratio μ/λ is small. Otherwise, if the required object is inhomogeneous, does not have a sharp boundary or contains a lot of noise and texture, a higher value of the ratio μ/λ is needed. Meanwhile, ν sets the penalty for the area inside the contour C . This parameter is essential only when two sides of boundaries (internal and external boundaries) are presented in the desired object (Getreuer, 2012). δ_o is a two-dimensional Dirac function that represents $\frac{d}{d\phi} H(\phi(x, y))$, ∇ (Eq. 2.75) is the gradient operator, and H is the Heaviside function (Chan and Vese, 2001; Klotz, 2013; Rousseau, 2009). Accordingly, by using level set function the Chan–Vese energy function can be rewritten as in Eq. 2.66 (Chan and Vese, 2001; Crandall, 2009; Hasan et al., 2016b) as follows.

$$\begin{aligned}
\mathcal{F}^{CV}(\phi(x, y)) &= \mu \int_{\Omega} \delta_o(\phi(x, y)) |\nabla\phi(x, y)| dx dy + v \int_{\Omega} H(\phi(x, y)) dx dy \\
&+ \lambda_1 \int_{\text{inside}(C)} |I(x, y) - c_1|^2 H(\phi(x, y)) dx dy \\
&+ \lambda_2 \int_{\text{outside}(C)} |I(x, y) - c_2|^2 (1 - H(\phi(x, y))) dx dy
\end{aligned} \tag{2.66}$$

The minimization is solved by alternatively updating c_1, c_2 and ϕ . Keeping ϕ fixed and minimizing the energy function \mathcal{F}^{CV} with respect to the optimal values c_1 and c_2 . Consequently, Eq. 2.67 and Eq. 2.68 are attained for c_1 and c_2 as functions of ϕ (Thapaliya et al., 2013; Chan and Vese, 2001; Tai et al., 2005; Hasan et al., 2016b).

$$c_1(\phi(x, y)) = \frac{\int_{\Omega} I(x, y) \cdot H(\phi(x, y)) dx dy}{\int_{\Omega} H(\phi(x, y)) dx dy} \tag{2.67}$$

$$c_2(\phi(x, y)) = \frac{\int_{\Omega} I(x, y) \cdot (1 - H(\phi(x, y))) dx dy}{\int_{\Omega} (1 - H(\phi(x, y))) dx dy} \tag{2.68}$$

To minimize the energy function \mathcal{F}^{CV} with respect to ϕ and fix the c_1 and c_2 , a gradient descent method is adopted and has yielded the associated Euler-Lagrange equation for ϕ , which is given by Eq. 2.69 (parameterizing the descent direction by an artificial time) (Thapaliya et al., 2013; Chan and Vese, 2001; Tai et al., 2005; Chan et al., 2000).

$$\begin{cases} \frac{\partial\phi}{\partial t} = \delta(\phi) \left[\mu \operatorname{div} \left(\frac{\nabla\phi}{|\nabla\phi|} \right) - v - \lambda_1 (I(x, y) - c_1)^2 + \lambda_2 (I(x, y) - c_2)^2 \right] & \text{in } \Omega \\ \frac{\delta(\phi)}{|\nabla\phi|} \frac{\partial\phi}{\partial \vec{n}} = 0 & \text{on } \partial\Omega \end{cases} \tag{2.69}$$

where \vec{n} represents the exterior normal to the boundary of $\partial\Omega$ and $\frac{\partial\phi}{\partial \vec{n}}$ represents the normal derivative of ϕ at the boundary.

2.6.4.4 3D Active Contour without Edge

The three-dimensional active contour without edge (3DACWE) method has the same principles of 2DACWE method that was explained in details in 2.6.4.3. The 3DACWE algorithm evolves the three-dimensional level set function and minimizes the Mumford-Shah functional. The initial contour C is defined as a 3D-box with x , y and z coordinates inside the desired object, such that x and y denotes the dimensions of 3D-box within each MRI slice and z denotes the number of MRI slices that are covered by 3D-box. Then the Chan-Vese energy function is given in Eq. 2.70 (Rousseau, 2009; Klotz, 2013; Hasan et al., 2016b).

$$\begin{aligned} \mathcal{F}^{CV}(C, c_1, c_2) = & \mu \text{Length}(C) + \nu \text{Area}(\text{inside}(C)) + \lambda_1 \int_{\text{inside}(C)} |I(x, y, z) - c_1|^2 dx dy dz \\ & + \lambda_2 \int_{\text{outside}(C)} |I(x, y, z) - c_2|^2 dx dy dz \end{aligned} \quad 2.70$$

where the first and second terms are given in Eq. 2.71 and Eq. 2.72 respectively (Chan and Vese, 2001; Klotz, 2013; Hasan et al., 2016b).

$$\text{length}(\phi(x, y, z)) = \int_{\Omega} \delta_o(\phi(x, y, z)) |\nabla \phi(x, y, z)| dx dy dz \quad 2.71$$

$$\text{area}(\phi(x, y, z)) = \int_{\Omega} H(\phi(x, y, z)) dx dy dz \quad 2.72$$

Accordingly, by using three-dimensional level set function the Chan–Vese energy function can be rewritten as in Eq. 2.73 (Chan and Vese, 2001; Crandall, 2009).

$$\begin{aligned}
\mathcal{F}^{CV}(\phi(x, y, z)) &= \mu \int_{\Omega} \delta_o(\phi(x, y, z)) |\nabla\phi(x, y, z)| dx dy dz + v \int_{\Omega} H(\phi(x, y, z)) dx dy dz \\
&+ \lambda_1 \int_{\text{inside}(C)} |I(x, y, z) - c_1|^2 H(\phi(x, y, z)) dx dy dz \\
&+ \lambda_2 \int_{\text{outside}(C)} |I(x, y, z) - c_2|^2 (1 - H(\phi(x, y, z))) dx dy dz \tag{2.73}
\end{aligned}$$

The minimization is solved by alternatively updating c_1 , c_2 and ϕ . Keeping ϕ fixed and minimizing the energy function \mathcal{F}^{CV} with respect to the optimal values of c_1 and c_2 . Consequently, Eq. 2.74 and Eq. 2.75 are attained for c_1 and c_2 as functions of ϕ (Rousseau, 2009; Klotz, 2013; Crandall, 2009; Hasan et al., 2016b).

$$c_1(\phi(x, y, z)) = \frac{\int_{\Omega} I(x, y, z) \cdot H(\phi(x, y, z)) dx dy dz}{\int_{\Omega} H(\phi(x, y, z)) dx dy dz} \tag{2.74}$$

$$c_2(\phi(x, y, z)) = \frac{\int_{\Omega} I(x, y, z) \cdot (1 - H(\phi(x, y, z))) dx dy dz}{\int_{\Omega} (1 - H(\phi(x, y, z))) dx dy dz} \tag{2.75}$$

To minimize the energy function \mathcal{F}^{CV} with respect to ϕ and fix c_1 and c_2 , a gradient descent method is adopted and has yielded the associated Euler-Lagrange equation for ϕ , which is given by Eq. 2.76 (parameterizing the descent direction by an artificial time) (Rousseau, 2009; Klotz, 2013; Crandall, 2009; Hasan et al., 2016b).

$$\begin{cases} \frac{\partial\phi}{\partial t} = \delta(\phi) \left[\mu \operatorname{div} \left(\frac{\nabla\phi}{|\nabla\phi|} \right) - v - \lambda_1 (I(x, y, z) - c_1)^2 + \lambda_2 (I(x, y, z) - c_2)^2 \right] & \text{in } \Omega \\ \frac{\delta(\phi)}{|\nabla\phi|} \frac{\partial\phi}{\partial n} = 0 & \text{on } \partial\Omega \end{cases} \tag{2.76}$$

where \vec{n} represents the exterior normal to the boundary of $\partial\Omega$, and $\frac{\partial\phi}{\partial n}$ represents the normal derivative of ϕ at the boundary.

To segment the brain tumour, it is essential to define an initial contour through tumour boundary. Then within iterations of 3DACWE, this contour evolves until reaching the actual border of the tumour. Several criteria can be used to stop the segmentation process; when the area of contour becomes constant or the energy function reaches the minimum value. The latter criteria will be sufficient to use in this study to stop the segmentation process.

2.7 Conclusion

In this chapter, a comprehensive literature on brain tumours analysis in MRI images was reviewed to select the most appropriate techniques and methods that were used for analysing MRI brain scans. The techniques were investigated based on various aspects such as feature extraction scheme and classification accuracy. The framework of this study includes image pre-processing, texture feature extraction, feature selection, classification and segmentation and a theoretical background for evaluating the selected methods were also demonstrated.

In the next chapter, the basics of magnetic resonance imaging and the characteristics of brain tumours will be introduced and explained in details.

CHAPTER THREE

Basics of Magnetic Resonance Imaging

Overview

This chapter includes an introduction to digital image representation, followed by fundamental principles of MRI with its important characteristics. Finally, a brief description about Al-Kadhimiya Teaching Hospital and some basic information about brain tumours are included.

3.1 Introduction

Image processing tools have been used significantly in medical imaging technologies and could improve the accuracy of the diagnostic processes. Medical imaging has been considered as a powerful technology used for clinical diagnosis. MRI has become one of the major research subjects in medical image processing field. Recently it has begun to be applied widely in screening of brain tumours, due to its sensitivity to locate changes in tissue density.

MRI is the standard important step for diagnosing and evaluating patients who have symptoms and signs of a brain tumour. It provides MR slices from different angles that enables clinicians to be more precise in diagnosing. In addition, it plays an important role in providing essential information such as tumour types (e.g. tumour or stroke), assessing tumour site, tumour area and volume, directing biopsies, planning the proper therapy and evaluating the therapeutic results (Drevelegas and Papanikolaou, 2011; Tonarelli, 2013; Mechtler, 2009).

3.2 Digital Image Representation

A digital image can be considered as a discrete representation of data which includes both spatial and intensity information. The digital image $I(m, n)$ can be represented by a series of numerical values of fixed positions ($m=1, 2, \dots, M; n=1, 2, \dots, N$) in a two dimensional array as shown in Fig. 3.1. The indices m and n specify the rows and columns of the image (Solomon and Breckon, 2011; El-Shenawy, 2013). A single position in a digital image is called picture element or pixel and is designated by (m, n) indices.

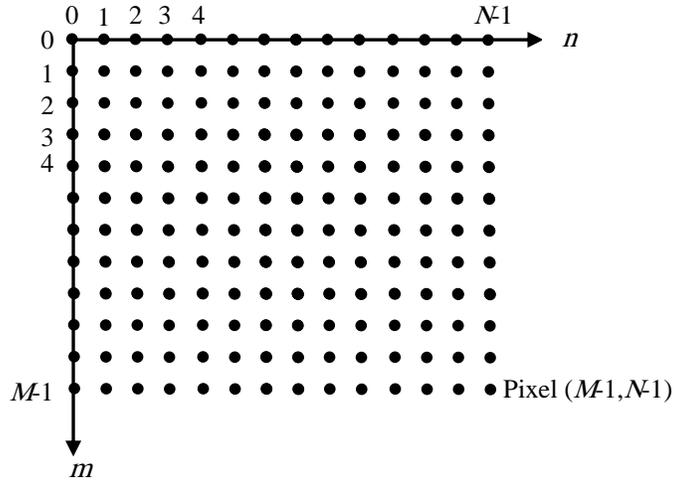


Figure 3.1: The representation of a digital image.

The associated numerical value refers to a grey value or colour information. In the case of grey scale image (8-bits), each pixel has a numerical value that is called intensity value and is typically in the range 0 to 255. Where, 0 represent the black colour and 255 represents the white colour.

In a three dimensional representation or volumetric data, a medical image is represented as $I(m, n, k)$, each picture element is called a voxel or volume element and represents the dimensions of the actual tissue. The indices m, n and k specify rows and columns of the medical image and the thickness of the voxel respectively (Sonka et al., 2014; Guy and Ffytche, 2005). The MRI scans includes a series of slices. Figure 3.2 shows an example of MRI brain scanning slices with the following MR sequences; the width is 512 pixel, the height is 512 pixel and the space between slices is 5.5 mm.

3.3 Magnetic Resonance Imaging

Medical imaging technology has experienced a dramatic change in the last 30 years (Christos, 2005). MRI is a vision and non-ionizing technology that has become most popular and widely used because of more precise and accurate imaging and diagnose of pathology in addition to excellent image contrast resolution. In contrast with other medical technologies such as X-ray and CT, it is harmless to the human body because of using radio frequency (200 MHz- 2 GHz) electromagnetic radiation and large magnetic field around (1-3 Tesla) to build cross sectional soft tissue visualization for all internal organs of the human body and vessels (Blink, 2004).

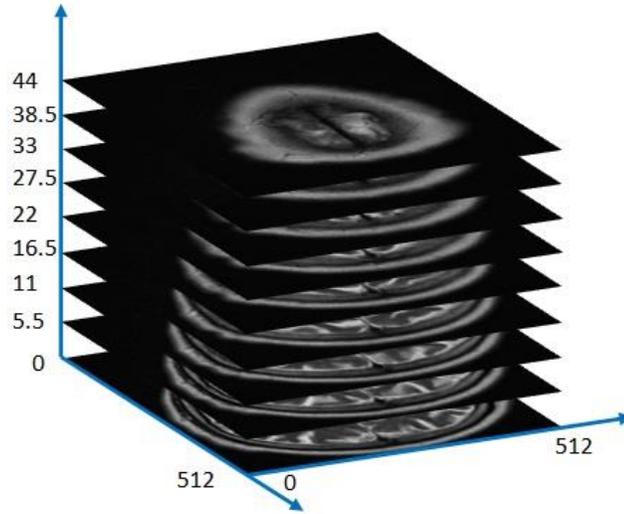


Figure 3.2: MRI brain slices with 5.5 mm thickness.

The MRI scanner has a powerful magnetic field which is generated by using a superconducting magnet. The wire has been used to construct the superconducting magnet with a resistance which approach to zero (Dougherty, 2009). The MRI has played an important role in the medical imaging and helped clinicians to diagnose and plan treatment such as surgery or radiation therapy. The majority of researches on medical imaging use MRI scans because it provides high resolution images, excellent soft tissue contrast and high signal to noise ratio (Pham et al., 1998; Blink, 2004). The MRI technology is based on the interaction between an external magnetic field and protons of hydrogen which possesses spin, since the human body consists of about 70% water. This mean that the biological tissue is composed of billions of hydrogen atoms. Therefore, the concentration is on the hydrogen nucleus which has a single proton (Petrou, 2011). The MRI is particularly well suited for the imaging of biological tissue such as brain, eyes and others rather than bones because they do not have many hydrogen atoms. The protons of the hydrogen are continuously moving. This movement is a self-rotation or spinning around an internal axis with a given value of angular momentum P . This rotation creates a magnetic field oriented with the direction of the internal axis of the rotation. Therefore, the proton has a magnetic moment μ as shown in Fig. 3.3. The magnitude of the angular momentum is a constant value and the magnetic moment of the proton can be calculated using Eq. 3.1 (Blink, 2004).

$$|\vec{\mu}| = \gamma |\vec{P}| \quad 3.1$$

where γ is a precession frequency of hydrogen protons; it is also known as the “gyromagnetic” ratio and it is equal to 42.576 MHz/Tesla.

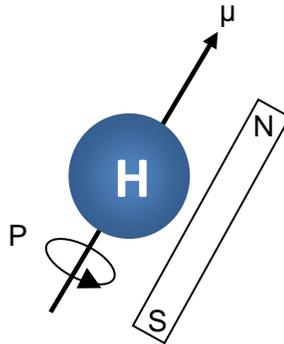


Figure 3.3: Internal rotation of a hydrogen proton.

This means that the hydrogen protons precess around the axis of external magnetic field about 42 million times per second (Schild, 1990). Normally, the direction of μ is randomly oriented in the absence of an external magnetic field as shown in Fig. 3.4, and it is aligned to an outer magnetic field if it is available (Petrou, 2011). Once applying a very strong external magnetic field on the human body, all the hydrogen protons align themselves along one of the two directions; parallel or anti-parallel to the external magnetic field B_o . Where, some of the protons align their magnetic moments with the same direction of B_o , they are called parallel protons. These protons have lower energy state. While the other protons align their magnetic moments in the opposite direction of the external magnetic field B_o and they are called anti-parallel protons. These protons have a higher energy state.

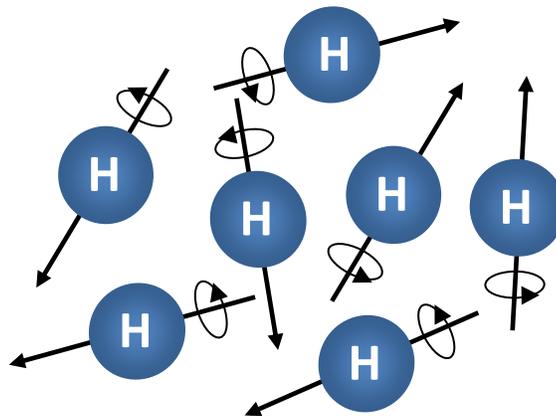


Figure 3.4: Random oriented poles of hydrogen protons.

Generally, the numbers of parallel protons are slightly more than anti-parallel protons as shown in Fig. 3.5. For instance, in 1.5 Tesla MRI scanner, for every 2 million protons, there are only 9 protons aligned along the direction of the external magnetic field B_o more than those which are aligned against it (Dougherty, 2009).

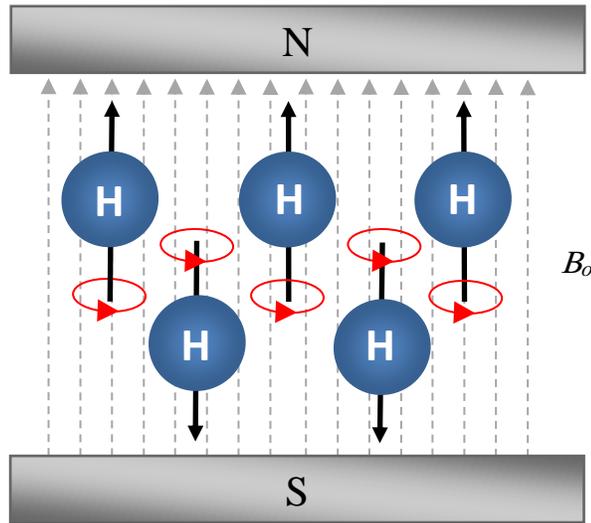


Figure 3.5: Aligned poles of hydrogen protons in external magnetic field.

Consequently, there is a net magnetization M_o which has the same direction of B_o as shown in Fig. 3.6, and created by finding the difference between the magnetic moments of parallel and anti-parallel protons by using Eq. 3.2 (Blink, 2004; Schild, 1990).

$$M_o = \sum_{Parallel} \mu - \sum_{Anti\ Parallel} \mu \quad 3.2$$

This means that the magnetic moment of the parallel protons cancels the magnetic moment of the anti-parallel protons. Precisely, these protons do not just orient their magnetic moment μ with the external magnetic field B_o but they start to spin around the axes of the external magnetic field B_o in a certain way as shown in Fig. 3.7. This process is called precession frequency or Larmor frequency. It is defined as the number of times that the protons precess per second.

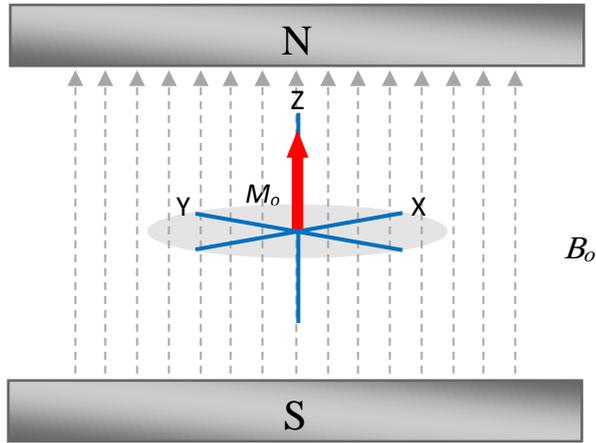


Figure 3.6: Direction of the net magnetization.

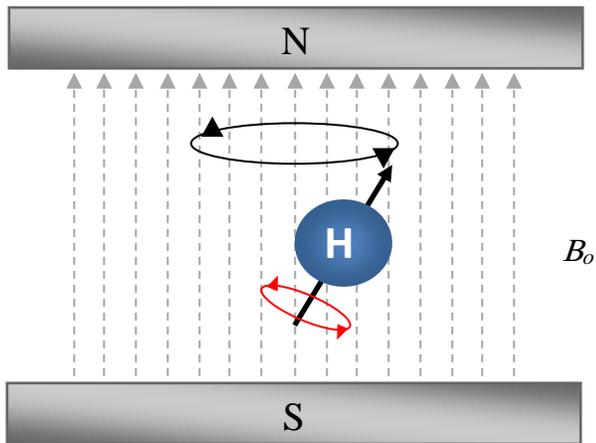


Figure 3.7: The precession movement of the proton.

It is not constant and depends on the intensity of the external magnetic field and is calculated using Eq. 3.3 (Blink, 2004; Schild, 1990).

$$\omega_o = \gamma B_o \quad 3.3$$

where ω_o is the Larmor frequency, B_o is the external magnetic field and γ is a precession frequency for the hydrogen proton.

In order to create detailed MRI slices, a second magnetic field is needed to excite the hydrogen protons and manipulate the net magnetization. An electromagnetic pulse, also known as a radio frequency (RF) pulse with frequency equal to precession frequency and perpendicular to the

direction of the external magnetic field B_0 . It is applied to incite resonance and causes some parallel protons to jump and to be in the higher energy anti-parallel direction (Dougherty, 2009). If 1.5 Tesla MRI scanner is used, RF pulse at frequency 63.855 MHz is applied and only protons that precess with the same frequency of RF pulse will respond to that RF pulse (Blink, 2004). Depending on time and amplitude of excitement; the net magnetization M_0 is effected and starts to tilt from the z-axis to the x-y plane as shown in Fig. 3.8, because the protons absorb energy from RF pulse and the number of high energy protons increases. Once the RF pulse is switched off, the excited net magnetization starts to relax slowly back to the original orientation along the z-axis which represent the system equilibrium state. This process is known as the relaxation process as shown in Fig. 3.9. The time that the net magnetization takes to return back to the original state includes essential information about the dynamics of the molecules in the patient (Petrou, 2011).

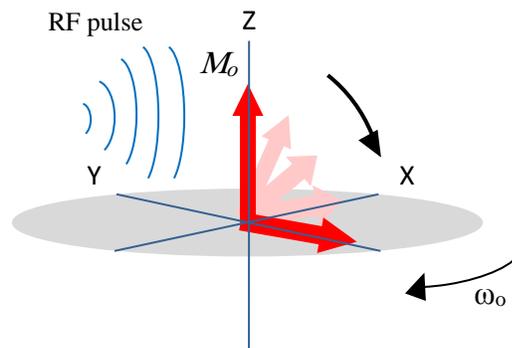


Figure 3.8: The net magnetization tilts toward the x-y plane.

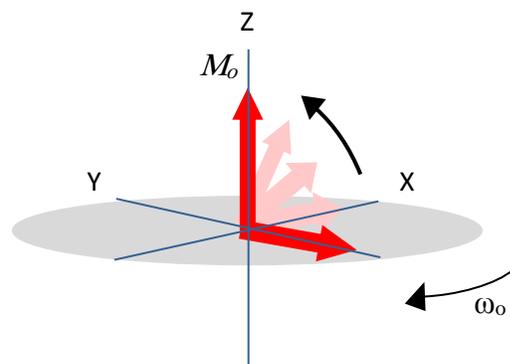


Figure 3.9: The net magnetization is relaxed after switching off the RF pulse.

There are two relaxation times that could be measured for protons namely T1 and T2 relaxation times; T1 relaxation is known as the *longitudinal* or *spin-lattice relaxation* that represents the required time for relaxing the excited net magnetization of protons back to recover 63% of the original net magnetization M_o after the RF pulse is switched off. The stored energy of protons is released as emitted signal during the relaxation process, which can be picked up by lattice (Petrou, 2011). Each tissue has different rate of relaxation process or energy releasing (e.g. T1 relaxation time is about 100 ms for fat and about 2000 ms for water) (Dougherty, 2009; Guy and Ffytche, 2005).

T2 relaxation is also known as *spin-spin relaxation* because it describes the interactions between the hydrogen protons in molecules (Blink, 2004). It represents the required time for declining the excited net magnetization M_{xy} to 37% of the original state due to all protons start to rotate at slightly different frequencies around the z-axis and start to exchange energy between each other (Dougherty, 2009; Blink, 2004; Guy and Ffytche, 2005).

T1 and T2 relaxations are two independent processes that happen simultaneously and are not correlated. Where, the T1 relaxation describes what happens in the z-axis, and the T2 relaxation describes what happens in the x-y plane. Additionally, the required time for T1 relaxation is always longer than T2 relaxation time. Different types of tissue have different relaxation time (e.g. fat tissue is de-phased quickly, while the water is de-phased much slower) (Blink, 2004). When both relaxation processes are terminated, the net magnetization returns to the original orientation along the direction of the external magnetic field B_o . Figure 3.10 shows the steps of excitation and relaxation of hydrogen protons in MRI. The hydrogen protons return to spin out of phase and lose the acquired energy as radio frequency waves. The RF coils that represent one of the most important components of MRI scanners, are used to transmit and collect the radio frequency waves. These waves are raw data, and picked up before disappearing in the space (Blink, 2004).

An MRI sequence, also known as the pulse sequence is defined as a sequence of events; RF pulses, gradient switches, data sampling periods and the timing between each of them are used to acquire the data from the human body and convert it to an image. The data is obtained by using a series of steps. First, switching on the slice selection gradient and simultaneously

applying an RF pulses to excite the net magnetization into the x-y plane. In addition, two essential elements of the sequence are the phase encoding and the frequency encoding which are required to localize the spatial resolution. Finally, the process is repeated many times by applying a series of excitation pulses, each separated by a repetition time (TR), the process of repetition is called gradient echo (GE) sequence.

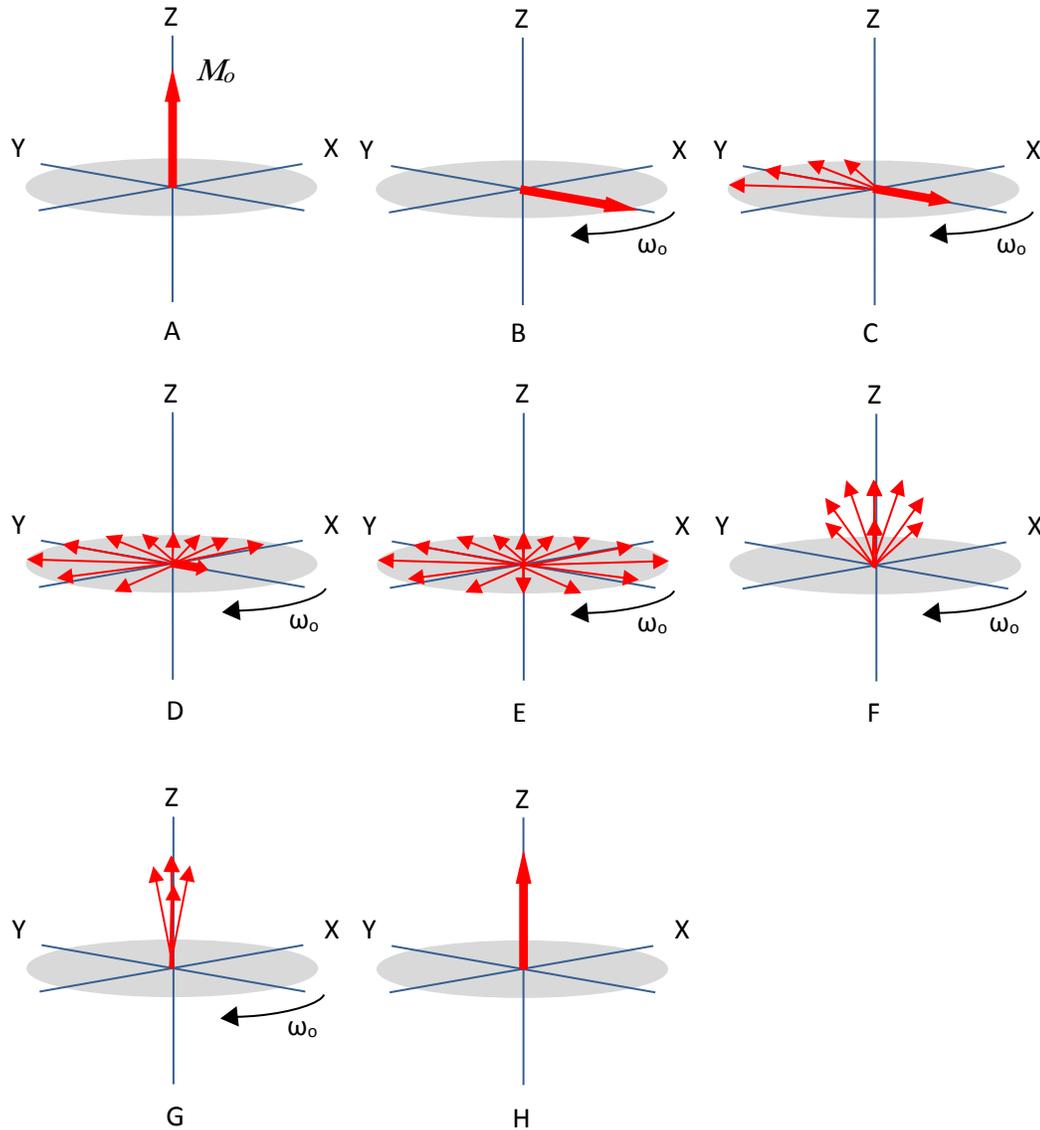


Figure 3.10: Excitation and relaxation of the hydrogen protons (Petrou, 2011), A) without excitation, B) with excitation, C,D,E) T2 relaxation (Spin-Spin Relaxation), F, G) T1 relaxation (Spin-Lattice Relaxation) and H) relaxation complete.

The TR is associated with T1 and T2 relaxations because T1 relaxation takes much longer than T2 relaxation therefore the TR should be short enough to allow T1 relaxation to fully complete and return to the equilibrium state. If there is not enough time for T1 to return to the equilibrium state, there will not be enough net magnetization available for the next repetition and this will lead to the loss of the signal (Blink, 2004). Additionally, collecting MRI signal requires an interval of time before applying a new excitation pulses and this time is called the echo time (TE). Changing TR and TE have an immediate effect on the contrast of the image (Nabizadeh, 2015).

To avoid the immediate dephasing process after switching off the RF pulse and to collect the best signal, a second 180° pulse is given, causes the spins to rephase. When all the spins are rephased, the signal to be higher again and ensure the acquisition of a much better signal. This process is called Spin Echo (SE) (Nabizadeh, 2015; Blink, 2004).

Inversion recovery (IR) sequence is a SE sequence preceded by another 180° invert excitation pulse (Blink, 2004). This inverts the initial longitudinal net magnetization of all tissues in the MRI slices to produce a heavily T1-w images to demonstrate anatomy and produce a large contrast difference between fat and water. Where, the full saturation of the fat and water vectors can be achieved when using the appropriate time to start IR (Nabizadeh, 2015).

Fluid attenuated inversion recovery (FLAIR) sequence is a special IR sequence that produces adaptive T2-w images by removing the signal of the brain edema and other structures with a high water content such as CSF (Mechtler, 2009). FLAIR is superior to T2-w images with respect to tumour delineation, better definition between edema and tumour, small hyper-intense tumours and the tumours that are adjacent to CSF (Kaal and Vecht, 2004; Nabizadeh, 2015).

3.4 Al-Kadhimiya Teaching Hospital

Al-Kadhimiya teaching hospital is one of the biggest hospitals in Iraq as shown in Fig. 3.11. It is located in Baghdad city and was established in 1984 with a capacity of 655 beds and then in 2000, it has been expanded to 812 beds. It provides care seven days a week, 24 hours a day. As well as there are 24 Consulting Clinics, each one receiving 100 patients daily. There are 2400 patients that are diagnosed daily. It includes three main centres:

- 1- Al-Jawad centre for tumours.
- 2- Hamed Al-Musaffa centre for dialysis.
- 3- Um Al-Baneen centre for in vitro fertilization and test tube baby.

The MRI Unit was established in the beginning of 2000 using a Philips Gyroscan 1.5 Tesla scanner. It received 25-35 patients daily and then it upgraded by installing a Siemens Avanto 1.5 Tesla scanner then followed by a Philips Achieva 1.5 Tesla scanner. Recently, it has been upgraded by installing a Philips Achieva 3.0 Tesla. Currently, the MRI unit provides services to more than 110 patients daily including brain test, abdomen test, spine test, etc.



Figure 3.11: Al-Kadhimiya Teaching Hospital.

3.5 Brain Tumours

The human nervous system is divided into two main parts; the central and the peripheral systems. The central nervous system is composed of the brain and the spinal cord. The peripheral nervous system is composed of the spinal nerves and the cranial nerves (Mayfield clinic for brain & Spine Institute, 2013). The human brain is divided into the left and right hemispheres that are approximately symmetric around the MSP.

Brain tumours are a heterogeneous group of neoplasms that vary significantly by morphological features, genetic alterations, growth potential, extent of invasiveness, tendency for progression or recurrence and treatment response (Mechtler, 2009). Brain tumours are abnormal and uncontrolled proliferations of cells. Some of these cells originate in the brain itself and they are called primary tumours. Others spread to the brain from somewhere else in the body through the blood stream and these are called secondary tumours. Primary brain tumours do not spread to other parts of the human body and they can be malignant or benign, while the secondary brain tumours are always malignant. The main characteristics of the benign tumour are that it grows very slowly and remain in the same part of the brain where it starts with clear defined boundaries. It does not destroy the surrounding area of the brain and does not spread to the other parts of the body. While malignant tumours represent life-threatening conditions because of their aggressive and invasive nature and their uncontrolled and fast growth. They have the ability to spread and damage the surrounding normal brain tissue (Pantelis, 2010). Both types of brain tumours are life threatening because the growth of brain tumours increases intracranial pressure inside the skull which has limited space and may cause edema, reduce blood flow and displacement with decaying of healthy tissue that controls vital functions (Tonarelli, 2013). Furthermore, stroke is also considered as one type of the brain abnormalities. It happens when the blood supply is cut off to some parts of the brain due to a blockage in the blood vessels. The strokes can affect patients in different ways depending on the part of the brain that has not received enough supply of the blood.

Generally, surgery is considered as the first step that should be taken to treat and achieve a gross total resection (GTR) of brain tumours. It is the preferable treatment when a tumour could be removed without any risks and side effects to the brain, in addition to reducing pressure that it generates inside the skull (American Brain Tumor Association, 2015). To determine the resection extent, MRI brain scanning should be used to calculate and compare the tumour volume on preoperative and postoperative MRI scan (Kim and Kim, 2012; American Brain Tumor Association, 2015). Generally, the MRI modalities can be categorized based on their advantages; T1-w images are not usually pathological but anatomical images and they are beneficial for black holes' detection which appear as hypo-intense or dark area relative to the WM intensities. On the other hand, T2-w images are sensitive to tissue pathology and show well-defined tumour delineation. They show the WM lesions as hyper-intense or bright area

relative to the WM intensities since most brain tumours are characterized by increased water content in tissues. Therefore, T2-w images are particularly useful for pathological detection (Tantisatirapong, 2015). The main drawback of this modality is that the CSF, GM and tumours have close intensities (Mortazavi et al., 2012). Clinically, T2-w and T1c-w are the mainstay of brain tumour diagnosis but using these two MRI modalities have sometimes difficulties in differentiating between the new and the old tumours or tumours from non-tumoural lesions in addition to grading (Tonarelli, 2013). Analysis of these diverse types of MRI images requires advanced computerized quantification and visualization tools and this could be supported by digital image processing technology. Image processing is concerned with the manipulation and interpretation of the objects in digital images by removing the effect of noise that is acquired during the acquisition process, deriving features and properties that could be used to discriminate the pathological patients and identify the location of abnormalities in the brain scans.

3.6 Conclusion

The representation of the digital image was explained throughout this chapter. The general basics of MRI scan was presented and described as well. Finally, the brain tumours section followed in this chapter was explained in detail.

In the next chapter, the pre-processing analysis of MRI brain scans techniques such as image enhancement and MSP detection and correction will be demonstrated and explained in details.

CHAPTER FOUR

Image Pre-processing Analysis of MRI

Overview

This chapter introduces different image pre-processing techniques that are used in this study to pre-process and enhance the MRI brain scans. It includes data collection, image resizing, image enhancement, intensity normalization and MSP detection and correction of the patient's brain.

4.1 Introduction

The main objective of this study is to develop and evaluate an automated screening system for MRI brain scans to discriminate normal and pathological patients without any clinician's interception; hence reducing clinicians' examination and interpretation time. The work will take place in five main stages; data collection phase from the Iraqi hospital, image pre-processing, feature extraction and classification, brain tumours location identification and segmentation. Figure 4.1 shows the overall flow chart of the proposed system.

The pre-processing step involves performing a set of algorithms on MRI brain scan slices as a preparation for the feature extraction step. This step includes resizing the dimensions of MRI slices, MRI slice enhancement by the Gaussian filter, intensity normalization of MRI slices and MSP of brain detection and correction algorithm. All the mentioned algorithms will be described in details with examples in this chapter.

4.2 Data Collection

Data collection is an essential step in this study and where two datasets are used. The clinical image dataset consists of 165 MRI brain scans acquired during routine diagnostic procedures at the MRI Unit of Al-Kadhimiya Teaching Hospital in Baghdad, Iraq. This dataset was diagnosed and classified into normal and abnormal by the clinicians of this unit. The MRI scans were obtained using a SIEMENS MAGNETOM Avanto 1.5 Tesla scanner (USA) and PHILIPS Achieva 1.5 Tesla scanner (Netherlands) as shown in Fig. 4.2.

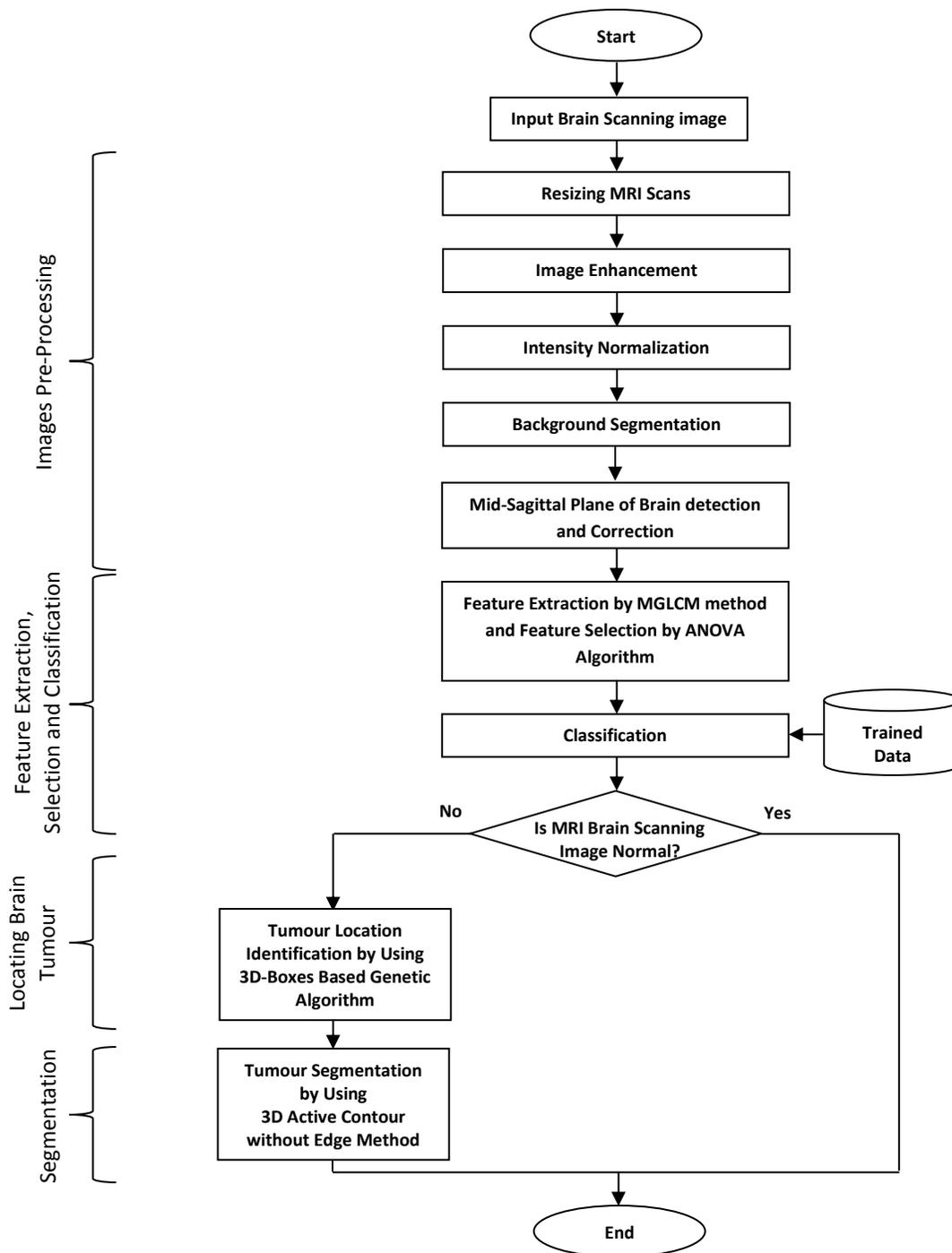


Figure 4.1: Overall flow chart of the proposed system.

The provided dataset consists of tumours with different sizes, shapes, locations, orientations and types. A total of 88 patients in this dataset exhibited different brain abnormalities with tumour sizes, shapes, locations, orientations and types.



Figure 4.2: SIEMENS MAGNETOM Avanto MRI scanner in MRI Unit of Al Kadhimiya Teaching Hospital- Iraq.

The remaining patients exhibited no detectable pathologies. The dataset included the four MRI image modalities, namely, T2-w, T1-w, T1c-w and FLAIR images under axial viewing and 1–5 mm slice thickness. 50 pathological patients in this dataset were manually segmented and labelled by an expert in this unit who evaluated the segmentation algorithm accuracy.

The standard benchmark multimodal brain tumour segmentation dataset (BRATS 2013) obtained from the international conference on medical image computing and computer-assisted interventions (Menze et al., 2015) was also used. This dataset includes 25 patients that were segmented manually by several human experts in addition to realistically generated synthetic brain tumours for which the ground truth segmentation is known. The synthetic MRI brain scans have a high variability in tumour shape and location, but they have less variability in intensity and less artefact-loaded than the real images to measure the capability of the algorithm to deal with variability in shape and the location of brain tumours. Few examples of the real and standard datasets that are used in this study with different modalities T2-w, FLAIR and T1c-w are shown in Fig. 4.3.

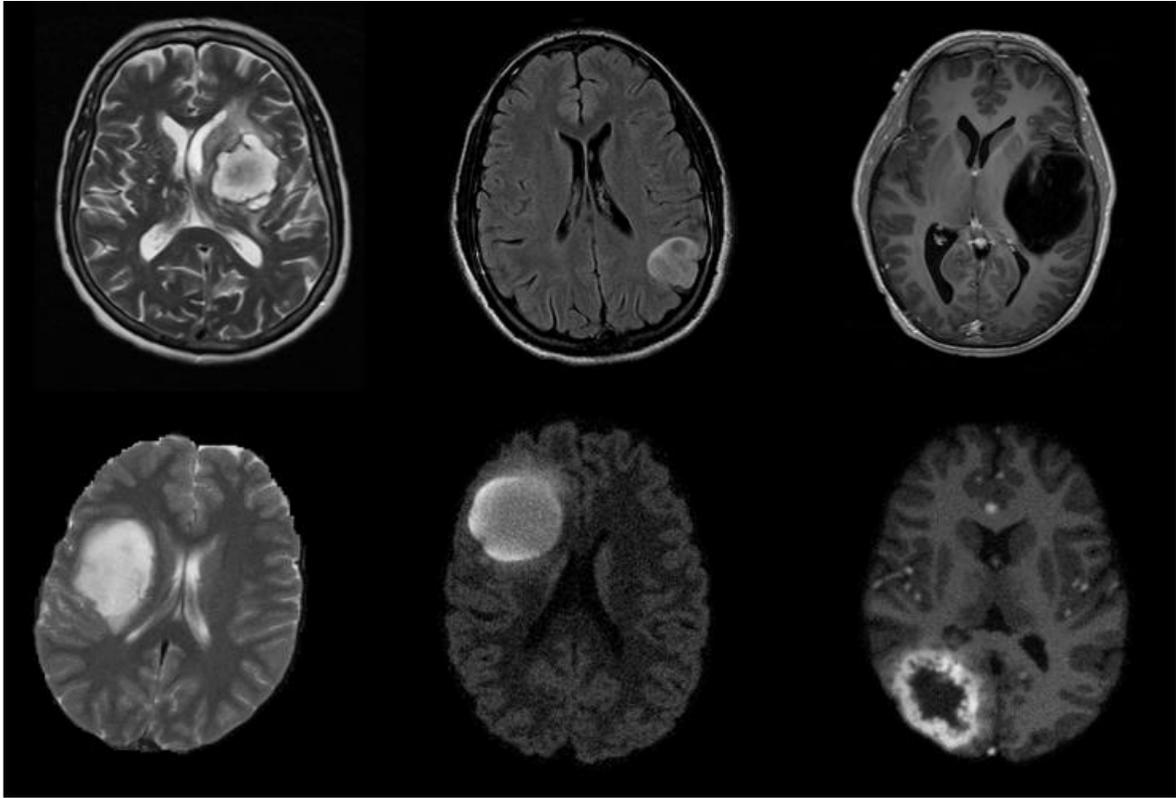


Figure 4.3: Examples from the used datasets: The collected dataset in the first row and the standard dataset in the second row.

4.3 Resizing the Dimensions of MRI slices

The provided MRI brain scans were collected from two scanners with different spatial resolutions. To enable the use of the full set without bias, the MRI scans were resized to 512×512 pixels. All algorithms developed in this study were implemented on squared slices. When the dimensions of the given MRI slices were changed to a square, care was taken to maintain the ratio of voxels to pixels (e.g., pixel spacing). Zero padding technique was used to adjust the MRI slice dimensions to become 512×512 pixels in resolution as shown in Fig. 4.4.

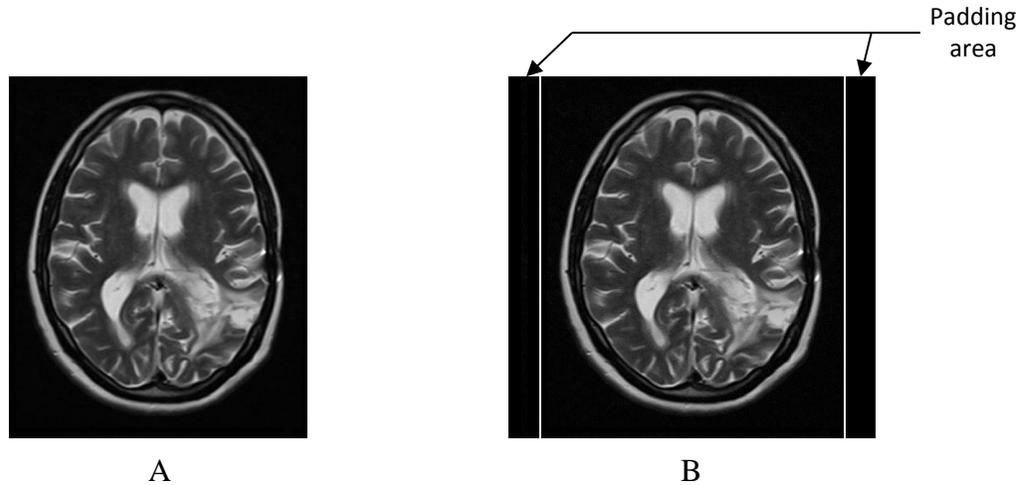


Figure 4.4: MRI slices resizing process, A) Original MRI slice with dimensions 512×470 , B) Resized MRI slice with dimensions 512×512 .

4.4 MRI Enhancement Algorithm

Image enhancement is a complex task that is highly dependent on the nature of the image. Several types of noise can be found in images and they require different image enhancement techniques. The typical noise in MRI slices appears as a small random modification of the intensity in an individual or small groups of pixels. These differences can be sufficiently large to lead to erroneous segmentation (Rousseau, 2009). The medical image visual quality plays an important role in the accuracy of clinical diagnosis because clinicians are usually trained and have experience with specific and high quality medical images. Generally, medical images are often contaminated by impulsive, additive or multiplicative noise during the acquisition and transmission processes. This leads to making the automatic feature extraction and analysis of clinical data a complicated task (Mohan et al., 2014; Pujar et al., 2010; William, 2001). As mentioned in chapter two, image enhancement algorithms are classified into spatial domain methods and frequency domain methods. A spatial domain low-pass filter (Gaussian filter) was applied for noise removal (Nabizadeh, 2015), such that the neighbourhood pixels are weighted according to the variance value σ in the Gaussian function (Birry, 2013; Gonzalez and Woods, 2002; Rogowska, 2009). The two dimensions Gaussian kernel operator $H(x, y)$ is given in Eq. 4.1 (Sonka et al., 2014).

$$H(x, y) = e^{\left(\frac{-(x^2+y^2)}{2\sigma^2}\right)} \quad 4.1$$

where x and y are the pixel coordinates in the image and σ is the standard deviation. The σ is proportional to the size of the neighbourhood pixels in the kernel operator which controls the degree of image smoothing. Consequently, a greater amount of smoothing is achieved when larger σ is selected (Birry, 2013; Nixon and Aguado, 2008; Schmid, 1999b; Rogowska, 2009). The results of applying Gaussian filters with different values of σ (0.5, 1 and 2) are shown in Fig. 4.5. This process is implemented by using *fspecial* and *imfilter* functions with Gaussian type in MATLAB R2013a Image Processing Toolkit (Matlab, 2013). The best value of σ was fixed to 0.5 experimentally.

4.5 Intensity Normalization

The pixel intensity values of each MRI slice were normalized to the same intensity interval to achieve dynamic range consistency. Histogram normalization was applied to stretch and shift the original histogram of the image and cover all the grayscale levels in the image using Eq. 2.1. The resulting normalized MRI slice achieved a higher contrast than that of the original slice because the histogram normalization method enhanced image contrast and provided a wider range of intensity transformation (Loizou et al., 2009; Tantisatirapong, 2015; Nabizadeh and Kubat, 2015; Sachdeva et al., 2012).

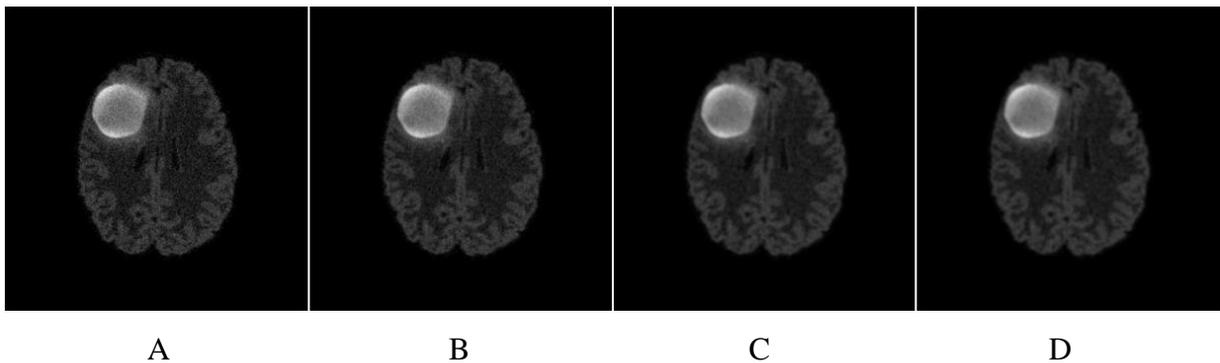


Figure 4.5: MRI brain scan image, A) Original image, B) Filtered image using Gaussian filter with ($\sigma=0.5$), C) Filtered image using Gaussian filter with ($\sigma=1$), D) Filtered image using Gaussian filter with ($\sigma=2$).

This approach demonstrated an enhanced classification of pathological tissues that can be achieved using the unmodified image (Tantisatirapong, 2015). Figure 4.6 shows the result of MRI slice normalization by the histogram normalization method.

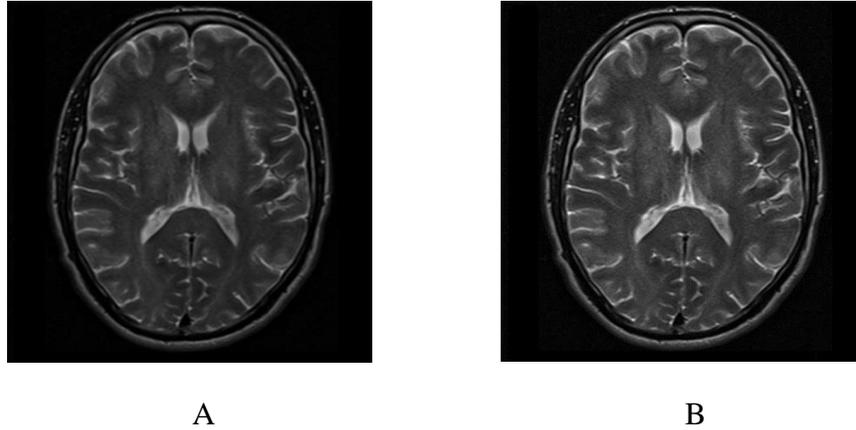


Figure 4.6: Histogram normalization method implementation, A) Original MRI slice, B) Normalized MRI slice.

4.6 Background Segmentation

Prior knowledge suggests that the background intensity values of MRI brain slices often approaches zero. The ability to eliminate and exclude the background of MRI brain slices is important because the background normally contains a much higher number of pixels than that of the brain region but without meaningful information (Liu et al., 1998; Nabizadeh, 2015). In this study, histogram thresholding was used as a segmentation method to isolate the background. This approach is based on the thresholding of intensity values by a specific T value. If the intensity value of a pixel is greater than T then the pixel is considered as a brain region, otherwise it is considered as a background (Naji et al., 2013). The T can be determined either manually which is specified by the user or automatically by using different approaches (Dougherty, 2009; Morris; Wilson and Ritter, 2000). Notably, the T2-w image histograms attained almost identical distribution shapes as shown in Fig. 4.7 (Udomchaiporn et al., 2013). Therefore, the T value was selected experimentally and set to 25 after the effects of a range of threshold values (13, 25, 50 and 76) were visually observed. This histogram thresholding is implemented by using *im2bw* function with specific value of threshold in MATLAB R2013a Image Processing Toolkit (Matlab, 2013).

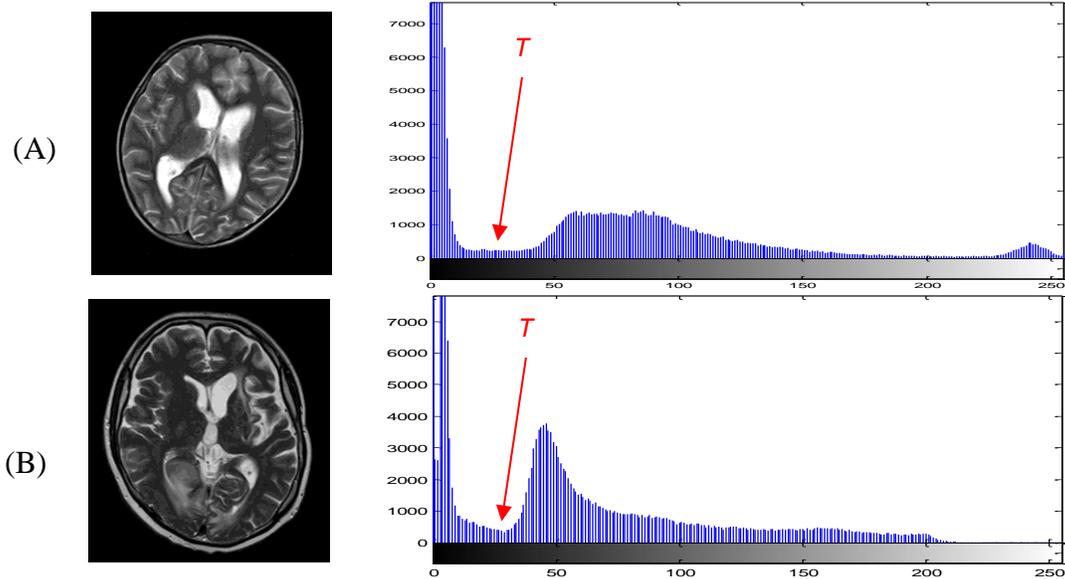


Figure 4.7: T2-w images of two different patients and corresponding histograms.

Subsequently, the application employs a set of morphological operators to remove any hole appearing in the region. There are many morphological operators but only two operators are essential and can be combined in many ways to produce more complex morphological operators which can solve different problems in image analysis. These two operators are dilation and erosion. The dilation is an operation that is used to increase the size of objects which are as foreground objects in binary images. While, the erosion is an operation that is used to increase the size of background and decrease the foreground objects in binary images (Dougherty, 2009; Bovik, 2009; Sonka et al., 2014). Additionally, holes filling morphological operator is used to fill holes that are defined as a background region of a binary image and surrounded by connected borders of foreground regions (Gonzalez and Woods, 2002; Soille, 2003; Wilson and Ritter, 2000). In this study, the deficiencies of the segmentation process are overcome by dilating the segmented MRI brain slice using the dilation morphological operator. Then the internal holes are filled using holes filling morphological operator. Consequently, a binary mask with ones denotes the patient's head, and zeros denotes the background. This mask is then multiplied with the original MRI brain slice to produce a new slice image without the background. Figure 4.8 shows an example of how an MRI slice is segmented, dilated and holes filled.

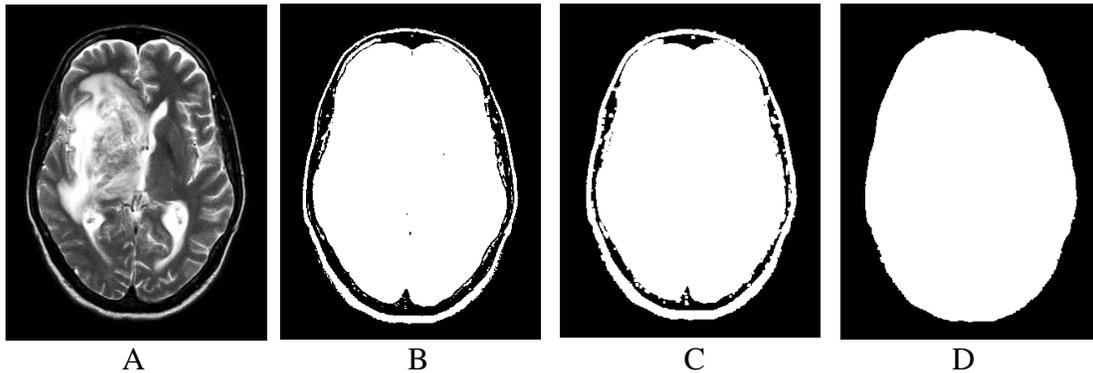


Figure 4.8: An example of skull boundary identification, A) Original MRI slice, B) Segmented MRI slice with threshold equal to 25, C) Dilated MRI slice and D) Filled holes in MRI slice.

4.7 MSP Detection and Correction

The advances in medical imaging techniques provide facilities for the internal visualization of the brain. These medical images are used for diagnosing and visual interpretation by clinicians. The MSP identification is an important initial step in brain image analysis because this method provides an initial estimation of the brain's pathology assessment and tumour detection (Jayasuriya and Liew, 2012). The human brain is divided into two hemispheres with an approximately bilateral symmetry around the MSP. This means that most structures in one side of the brain have a counterpart on the other side with similar shape and location. The two hemispheres are separated by a longitudinal fissure that represents a membrane between the left and right hemispheres. This longitudinal fissure is filled with CSF and it can be used to recognize the two hemispheres visually (Ruppert et al., 2011). The two hemispheres separation process in the axial MRI brain slices can be done by recognizing the MSP along the longitudinal fissure which can be used as a reference for asymmetry analysis. The MSP of the brain has the same orientation of the patient's head. The symmetry of the brain is an important indicator about its normality or abnormality such that most pathologies such as tumours, bleeding and stroke can be determined by a symmetry based analysis of MRI brain scan. However, the growth of tumour cells can destroy the symmetry and curve the MSP of the brain (Liu et al., 1998). The MSP extraction methods can be divided into two groups: content-based methods and shaped-based methods (Ruppert et al., 2011; Liu, 2009).

The content-based methods are based on finding a plane that maximizes a symmetrical measure between both sides of the brain (Christensen et al., 2006; Ardekani et al., 1997; Khotanlou et al., 2009; Ruppert et al., 2011; Liu, 2009; Schmid, 1999a). The major obstacle preventing these methods from a wide adoption in realistic neuro-application is the difficulty of measuring symmetry and identifying the MSP of the brain for the pathological cases e.g. the air pockets and the presence of lesions should be ignored when computing the axis of symmetry (Liu et al., 1998; Hu and Nowinski, 2003). By contrast, shaped-based methods use the inter-hemispheric fissure as a simple landmark to extract and detect the MSP of the brain which denotes the symmetry plane (Bergo et al., 2008; Liu, 2009). All parallel axial slices, the inter-hemispheric fissure lines are parallel with the same orientation of patient's head (Hu and Nowinski, 2003). In this study, we focused on determining the orientation of the patient's head instead of measuring the symmetry to identify the MSP of the brain (Hasan and Meziane, 2016). The proposed method is based essentially on using the PCA method to compute the distinctive principle axes that are orthogonal to each other. Those axes are used to characterize the patient's head by representing the spatial distribution of the mass. Where, any plane of symmetry in the body is orthogonal to a principle axis (Liu, 2009; Hasan and Meziane, 2016).

4.7.1 Orientation Determination Based PCA

Several methods have been proposed to determine the orientation of objects in images. The most widely used method is PCA (Schmid, 1999b) that was explained in details in section 2.4.1.2. The PCA method essentially attempts to transfer the coordinates of the original data to a new coordinate system such that the maximum variation in the data comes to lie on the first coordinate. This is known as the first principal component. The second maximum variation in the data lies on the second coordinate and so on. The most common steps that are followed by radiologists and clinicians in MRI units and specifically in MRI Unit of Al-Kadhimiya Teaching Hospital, include positioning and aligning the patient's head inside the head coil according to the laser light indicator as shown in Fig. 4.9, and using sponges to support and minimize the head tilt and rotation. This gives better MRI image quality (UC Davis Medical Center, 2016).



Figure 4.9: How the patient's head lies in the MRI head coil
(UC Davis Medical Center, 2016).

Due to all brain slices in the same scan having the same symmetry axis orientation (Liu et al., 1998), it is possible to detect the degree of skewness to the left or right by using single slice in axial viewing instead of using all brain slices in the context of reducing computational complexity. In this study, we assume that the patient's head may be skewed only either left or right. Let D be an original two-dimensional data with two observations that are plotted on X and Y coordinates. The PCA is used to map linearly these coordinates into new X' and Y' coordinates, where X' extends along the direction of the maximum variation of a given data and Y' is perpendicular to X' and extends along the direction of the minimum variation of a given data as shown in Fig. 4.10 (Wallisch et al., 2014).

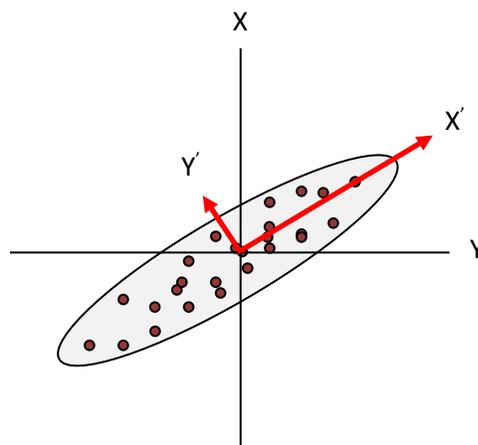


Figure 4.10: Remapping the axes (X, Y) of the original data into new axes (X', Y') .

In this study, D represents the coordinates of pixels in the foreground part of the segmented MRI brain slice, such that $X = [X_1, X_2, \dots, X_n]$, and $Y = [Y_1, Y_2, \dots, Y_n]$. These coordinates are normalized by subtracting the mean from each one according to Eq. 4.3 and Eq. 4.4.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad 4.3$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad 4.4$$

The covariance matrix (*cov*) that is a symmetrical and a semi-positive definite matrix, is used to measure to which extent these coordinates are linearly related and is given in Eq. 4.5.

$$cov(x, y) = \frac{1}{(n-1)} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})^T \quad 4.5$$

If the given data has m dimensions, the covariance matrix is an m by m matrix (Wallisch et al., 2014; Manly, 1994). Then the eigenvectors and eigenvalues can be calculated by using Eq. 4.6 and Eq. 4.7 respectively.

$$|cov - \lambda I| = 0 \quad 4.6$$

$$cov.V = \lambda V \quad 4.7$$

where λ is the eigenvalues of the covariance matrix, I is the identity matrix and V is the eigenvectors matrix.

The eigenvectors and eigenvalues include useful information about the new coordinates of the given data (Smith, 2002). Each eigenvector points in the direction of a new coordinate axis. The desirable coordinate that has the highest eigenvalues and passing through the maximum variation of data, represents the orientation of the patient's head (Wallisch et al., 2014; Sonka et al., 2014). Then, the angle θ between the X-axis and X' -axis represents the degree of skewness of patient's head during the MRI test as shown in Fig. 4.11 and is calculated using Eq. 4.8.

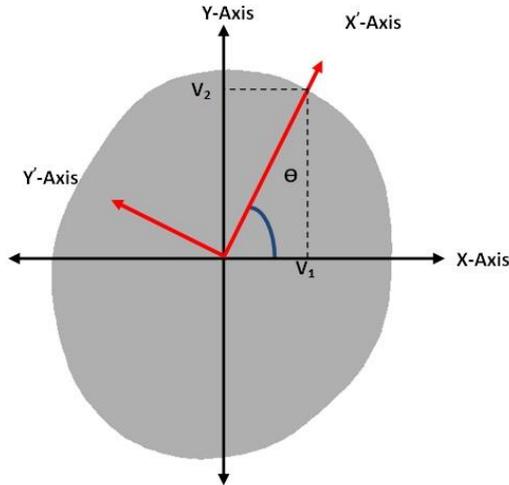


Figure 4.11: Original and new coordinates of brain.

$$\theta = \arctan\left(\frac{V_2}{V_1}\right) \quad 4.8$$

where V_1 and V_2 are the eigenvectors which are related to the maximum eigenvalues.

The main shortfall of PCA that it is not efficient in distinguishing between the axis of symmetry and axis of orientation. However, it is still an interesting approach because of its simplicity and the low processing time (Schmid, 1999a). The PCA algorithm is implemented using *princomp* function in MATLAB R2013a Image Processing Toolkit (Matlab, 2013).

4.7.2 Geometrical Transformation of Patient's Head

Geometrical transformation methods are widely used in computer graphic and image analysis. They help to eliminate the geometric distortion that occurs within image capturing (Sonka et al., 2014). They can be used to estimate the unknown pixels by the interpolation of the input pixels and rotating the object around a fixed point known as the centre of rotation (William, 2001). A geometric transformation includes two basic steps. First, the pixel coordinates transformation and second, the brightness interpolation (Sonka et al., 2014). There are two types of interpolation methods; nearest neighbour interpolation and bilinear interpolation (Wilson and Ritter, 2000; William, 2001).

A- Nearest Neighbour Interpolation

Nearest neighbour interpolation method is used to rotate an image through a given angle θ by using Eq. 4.9, Eq. 4.10 and Eq. 4.11 respectively. Where, the pixel with x and y coordinates in the original image is mapped into x' and y' coordinates in the resultant image by interpolating an output pixel location between four input pixels (Wilson and Ritter, 2000; William, 2001).

$$x' = x \cos \theta - y \sin \theta \quad 4.9$$

$$y' = x \sin \theta + y \cos \theta \quad 4.10$$

$$f'(x, y) = f([x'], [y']) \quad 4.11$$

where f' is the new interpolated image and $[]$ denotes the rounding of the new mapped coordinates. The main problem of this method is the probability of mapping the positions of the input pixels with integral coordinates to non-integral coordinates positions in the output image. The new pixel locations are generally in somewhere between four neighbouring pixels in the given image as shown in Fig. 4.12.

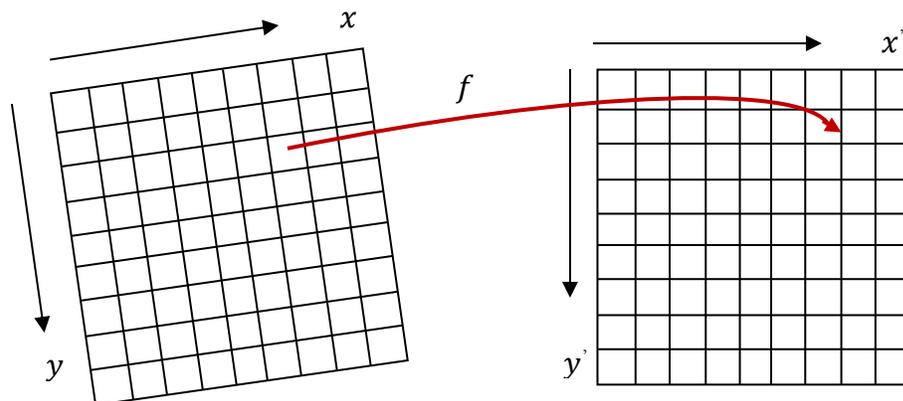


Figure 4.12: Mapping pixels by nearest neighbour interpolation.

B- Bilinear Interpolation

Bilinear Interpolation method is used to rotate an image by locating the new pixel somewhere between four neighbouring pixels of a given image as shown in Fig. 4.13, and using Eq. 4.12.

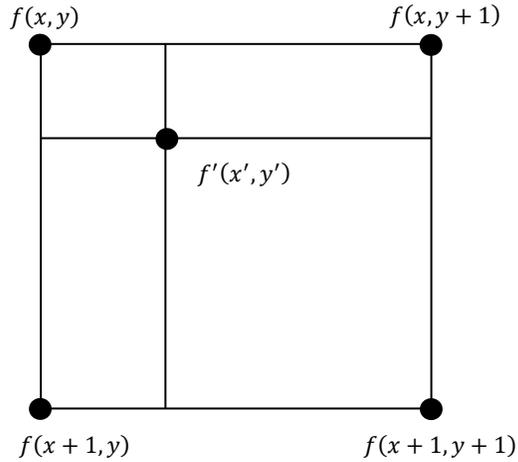


Figure 4.13: Mapping pixels by bilinear interpolation.

$$\begin{aligned}
 f'(x', y') &= f(x, y) + f(x + 1, y)(x' - x) - f(x, y)(x' - x) + f(x, y + 1)(y' - y) \\
 &\quad - f(x, y)(y' - y) \\
 &\quad + [f(x + 1, y + 1) + f(x, y) - f(x, y + 1) - f(x + 1, y)] (x' - x)(y' \\
 &\quad - y)
 \end{aligned} \tag{4.12}$$

It is a more desirable method but it has a problem with computational complexity (Wilson and Ritter, 2000). It is used in this study to rotate and correct the patient's head after the degree of wobbling θ is calculated and it is implemented using *Geometric Rotator system object* in MATLAB R2013a Image Processing Toolkit (Matlab, 2013).

4.7.3 Centralize Patient's Head in the Centre of MRI Slice

The patient's head is positioned in the centre of the MRI slice because of identifying the brain's abnormality depends essentially on measuring the symmetry between the two brain's hemispheres and the centroid of the patient's head is identical with the MSP of the brain (Liu et al., 1998). Therefore, it becomes easy to make the MSP of the brain exactly in the centre of the MRI brain slice by shifting the patient's head either left or right using Eq.4.13 and Eq. 4.14.

$$g_x = \frac{1}{N} \sum_{i=1}^N x_i \quad 4.13$$

$$g_y = \frac{1}{N} \sum_{i=1}^N y_i \quad 4.14$$

where N is the number of pixels within the segmented patient's head, and g_x and g_y are the coordinates of the centroid. Then the patient's head is shifted by a number of pixels that is equal to the difference between g_y and 256, which represent the coordinates of the middle line of the MRI brain slice.

Since the MRI brain slices of each patient have the same MSP orientation (Liu and Collins, 1996), the MSP detection and correction algorithm is implemented on a single slice instead of using all the slices to avoid computational complexity. The preferable slice for implementing the MSP detection and correction algorithm is the slice which locates in the lower of the brain and contains the largest number of pixels. It provides more accurate detection rate compared to slices higher in the brain (at the tip of the head) which have ovals or near-circular shape (Liu et al., 1998). Figure 4.14 shows the result of the MSP detecting and correcting of the three MRI brain slices which are shown in different orientation in the first column on the left of the figure. In the fifth column on the right of the same figure, the MRI slices are corrected and aligned in the middle of the slice.

To compare with an expert clinicians' delineation, the MSPs of 50 MRI scans from the collected dataset were manually identified by expert clinicians from the MRI Unit in Al-Kadhimiya Teaching Hospital. These MRI scans were given to the experts after correcting and aligning the MSPs of these scans. The proper location of the fitted line was drawn with the computer mouse by the experts. Figure 4.15 shows the results of delineation of MSPs of three MRI slices by the expert clinicians and our algorithm. Figure 4.16 shows the mean squared error (MSE) distribution between manual and our algorithm delineation of MSPs. Consequently, 86% of the computed MSPs are matched approximately with the clinicians' delineation within $MSE \leq 3^\circ$.

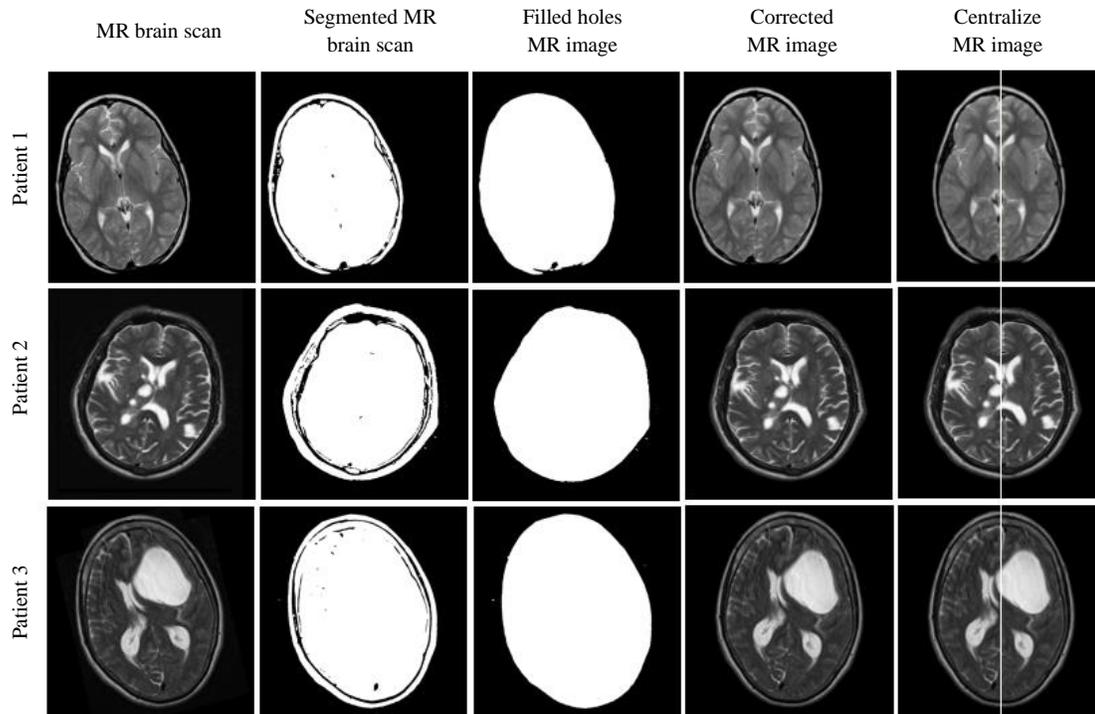


Figure 4.14: Steps for detecting and correcting the MSP of the brain in MR brain slices.

For further evaluation, the MRI brain slices shown in Fig. 4.17, are re-sampled using the Geometric Rotator system object in MATLAB R2013a Image Processing Toolkit (Matlab, 2013), to rotate the patient's head with yaw angles from -10 to 10 degrees in 2.5 degree intervals.

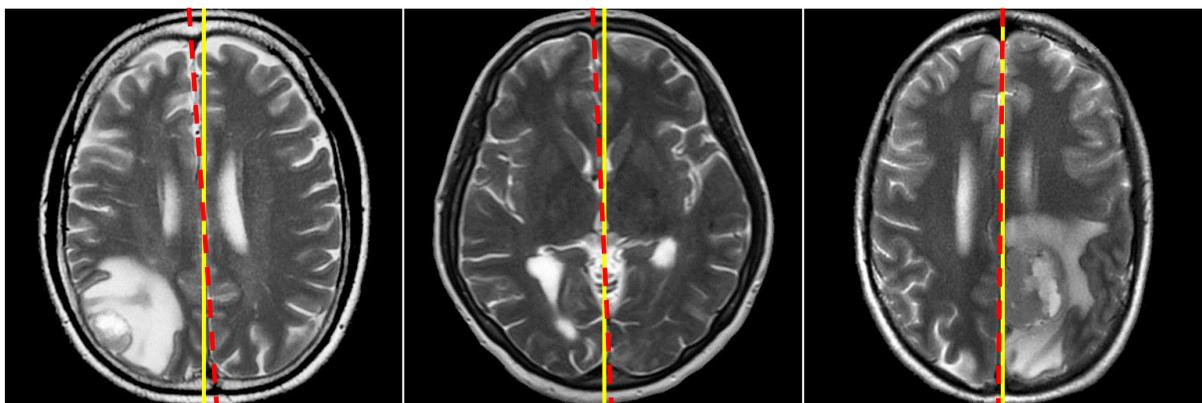


Figure 4.15: The solid yellow lines denote the computed MSPs and the dashed red lines denote the clinicians' delineation.

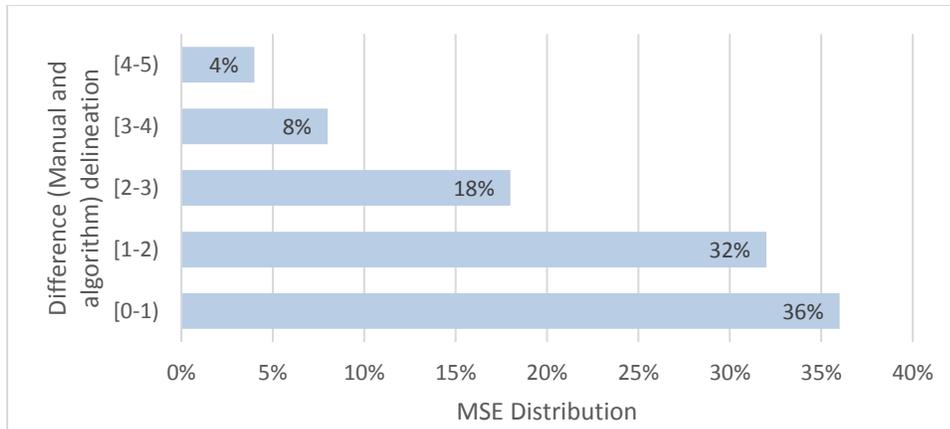


Figure 4.16: Distribution of MSE between manual and our algorithm delineation of MSPs.

Table 4.1 shows the experimental results of the predicted yaw angles of patient's head with the average MSE value. It seems that the predicted yaw angles approach to the actual yaw angles of patient's head.

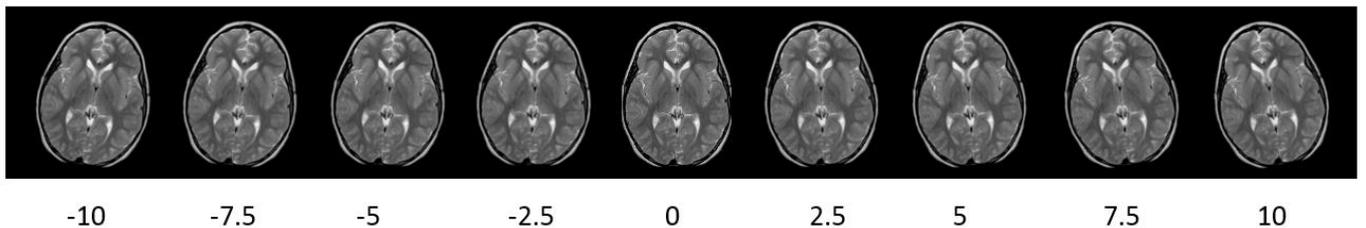


Figure 4.17: Resampling of one slice from the axial MRI brain scanning image with varied rotate angles.

Table 4.1: The results of predicted yaw angle.

Yaw Angle	-10°	-7.5°	-5°	-2.5°	0°	2.5°	5°	7.5°	10°	MSE
Proposed Algorithm	-9.08°	-7°	-4.74°	-2.43°	0.58°	3.3°	5.5°	8.21°	10.57°	0.35°

4.8 Conclusion

In this chapter, we described how the datasets were collected and unified to dimensions 512×512 . This was followed by the description of four main pre-processing algorithms; MRI enhancement, intensity normalisation, background segmentation and MSP detection and correction. The MSP of the brain detection and correction algorithm can automatically locate the MSP of given MRI brain scans. The algorithm works on both normal and pathological brain scans. These algorithms were used to prepare the MRI brain scans for next step of texture feature extraction and classification that will be explained in details in the next chapter.

CHAPTER FIVE

Features Extraction and Classification

Overview

This chapter covers the implementation of the proposed MGLCM method for texture feature extraction in addition to some prior processing steps that should be implemented primarily to prepare the MRI brain scans. It includes also the implementation of ANOVA for selecting the most relevant predictors and different techniques for classification such as LDA, SVM and ANN as an intelligent classifier.

5.1 Introduction

The fundamental objective of any diagnostic medical imaging investigation is tissue characterization. The texture analysis is an important way to provide unique information on the texture or spatial variation of pixels from medical images (Nabizadeh, 2015). Texture analysis methods are useful for studying and discriminating between pathologically different regions on medical images. It provides better performance than human eyesight in discriminating certain classes of texture. Practically, it requires careful consideration of the significance of the individual features to achieve high discrimination by reducing the effect of heavily correlated features and the features with little discriminatory power (Nailon, 2010). Texture refers to properties that represent the surface or structure of an object such as smoothness, coarseness and regularity. These properties are used to quantify the texture content of an object. Texture can be defined as an existing relationships of related pixels and group of pixels, this group of pixels is known as texture primitives or texture elements (Sonka et al., 2014; Nabizadeh, 2015).

In this study, only the textural features are considered because there is no colour information or regular shapes of the tumours that could be extracted, in addition they may appear in different image intensities (Prastawa et al., 2004). The texture features will be extracted from MRI brain slices to encode clinically valuable information using the proposed modified grey level co-occurrence matrix (MGLCM) method. The texture features will also be used to measure

statistically the similarity between the two separated hemispheres of the brain. A prior pre-processing algorithms that should be taken to prepare the MRI brain slices for texture features extraction by the proposed method as shown in Fig. 5.1.

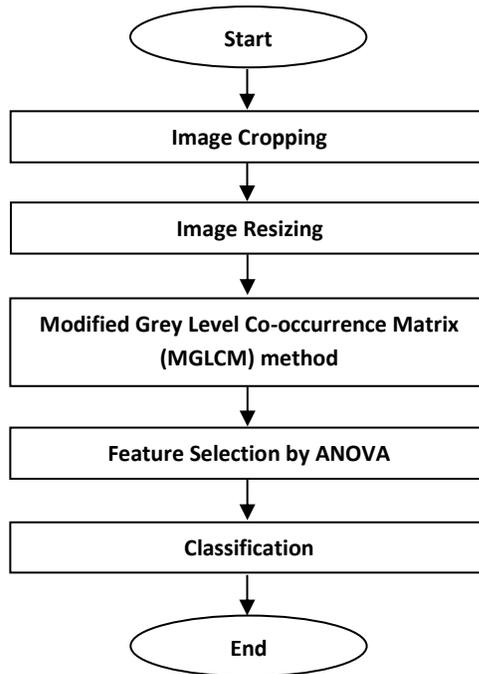


Figure 5.1: Flowchart of implementing feature extraction, selection and classification.

5.2 Preparing MRI Brain Slices for Feature Extraction

This step includes the implementation of a set of image pre-processing algorithms to prepare and make the MRI brain slices more appropriate for implementing the MGLCM method. The input for this step is a corrected MRI brain scans and the output includes only the patient's head with dimensions of (512×512) pixels.

5.2.1 Image Cropping

The MRI slices are cropped from the upper margin of the slices to the upper boundary of the skull. The same procedure is then used for cropping the MRI slices from the bottom margin of the slices to the bottom boundary of the skull. The left and right boundaries can be identified

by measuring the distance between the upper and bottom boundaries. In fact, the left and right boundaries are away from the middle of the MRI slices which denotes the brain MSP by a half distance between the top and bottom boundaries as shown in Fig. 5.2.

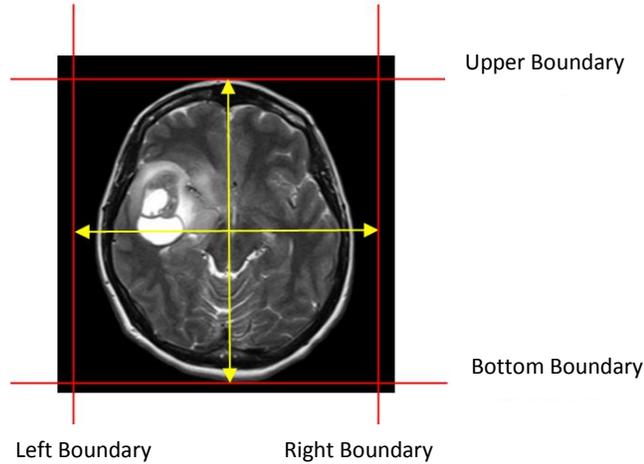


Figure 5.2: MRI brain slice cropping.

5.2.2 Image Resizing

Image resizing or specifically image zooming is an important process that is used in a variety of applications where specific number of pixels are inserted between the actual pixels of MRI brain slice to expand its size. The intensity values of the new pixels are interpolated from the surrounding original pixels (Bovik, 2009). In this study, every cropped MRI brain slice is resized to (512×512) pixels before using the MGLCM method. The image resizing algorithm is implemented by using *imresize* function in MATLAB R2013a Image Processing Toolkit (Matlab, 2013).

5.2.3 Modified Grey Level Co-occurrence Matrix Method

Modified grey level co-occurrence matrix (MGLCM) is a second-order statistical method proposed by Hasan and Meziane (2016) to generate textural features and provide information about the patterning of MRI brain scans textures. These features are used to measure statistically the degree of symmetry between the two brain hemispheres. Symmetry is an important indicator that can be used to detect the normality and abnormality of the human brain. MGLCM generates texture features by computing the spatial relationship of the joint frequencies of all pairwise

combinations of grey-level configuration of each pixel in the left hemisphere which is considered as a reference pixel, with one of nine opposite pixels existing in the right hemisphere under nine offsets $\theta = (45,45), (0,45), (315,45), (45,0), (0,0), (315,0), (45,315), (0,315), (315,315)$, and one distance $d=1$, as shown in Fig. 5.3. Consequently, because each pixel on the left hemisphere has nine opposite pixels on the right hemisphere, nine co-occurrence matrices are obtained for each MRI brain scanning image.

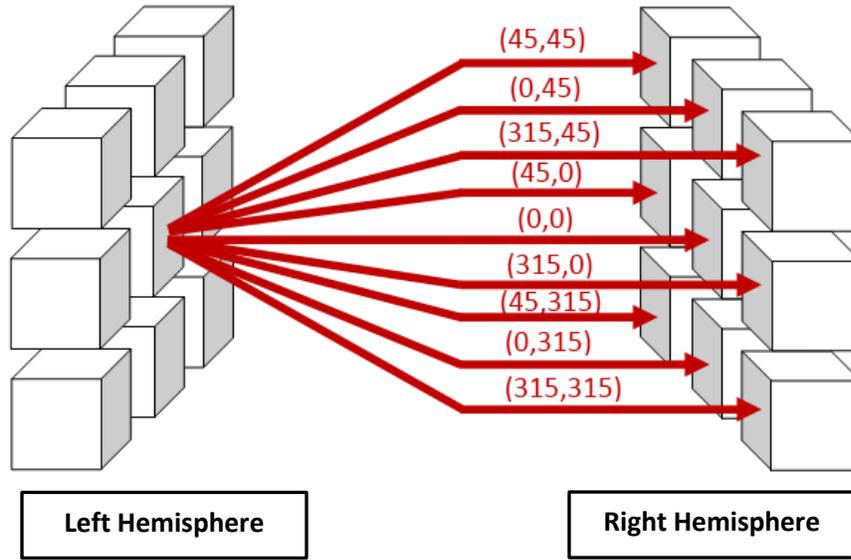


Figure 5.3: How reference pixel relates with opposite nine pixels.

Thereafter, each co-occurrence matrix is normalized by the total number of all its elements to calculate the co-occurrence relative frequency between the grey levels of joint pixels in the brain hemispheres. The nine co-occurrence matrices are defined by Eq. 5.1.

$$P(i,j)_{(\theta_1,\theta_2)} = \frac{1}{256^2} \sum_{x=1}^{512} \sum_{y=1}^{256} \begin{cases} 1, & \text{if } L(x,y) = i \text{ and } R(x + \Delta x, y + \Delta y) = j \\ 0, & \text{Otherwise} \end{cases} \quad 5.1$$

where L and R denote the left and right hemispheres respectively and both of them have a size of (512×256) pixels. P is the resultant co-occurrence matrix. i and j are the coordinates of the co-occurrence matrix. Δx and Δy values depend upon the directions of measured matrix and are obtain using the following rules:

If $\theta_1=0$ and $\theta_2=0$ then $\Delta x=0$ and $\Delta y=0$,
 If $\theta_1=0$ and $\theta_2=45$ then $\Delta x=-1$ and $\Delta y=0$
 If $\theta_1=0$ and $\theta_2=315$ then $\Delta x=1$ and $\Delta y=0$,
 If $\theta_1=45$ and $\theta_2=0$ then $\Delta x=0$ and $\Delta y=1$,
 If $\theta_1=315$ and $\theta_2=0$ then $\Delta x=0$ and $\Delta y=-1$,
 If $\theta_1=45$ and $\theta_2=45$ then $\Delta x=-1$ and $\Delta y=1$,
 If $\theta_1=315$ and $\theta_2=45$ then $\Delta x=-1$ and $\Delta y=-1$,
 If $\theta_1=315$ and $\theta_2=315$ then $\Delta x=1$ and $\Delta y=-1$,
 If $\theta_1=45$ and $\theta_2=315$ then $\Delta x=1$ and $\Delta y=1$.

The resultant co-occurrence matrices are approximately symmetric around the forward diagonal of the matrix for a healthy brain and asymmetrical for pathological patients. Figure 5.4 shows two examples of normal and abnormal MRI brain scans and the corresponding co-occurrence matrix at angles $\theta_1=0$ and $\theta_2=0$. On the left is the MRI scan and corresponding co-occurrence matrix of a normal brain scan. The MRI scan of normal patient shows that the hemispheres of brain are approximately symmetry around the MSP and the corresponding co-occurrence matrix is slightly narrower and symmetry around the forward diagonal. While on the right is the MRI scan and corresponding co-occurrence matrix of an abnormal brain scan. The patient has a tumour in left hemisphere of his brain and it makes his brain asymmetry around MSP. Again, the corresponding co-occurrence matrix is significantly wider and asymmetry around the forward diagonal.

To reduce the dimensionality of the feature space, the resultants MGLCM matrices of all the MRI slices at all orientations were added. The maximum number of grey levels considered for each slice was typically scaled down to 256 grey levels (8 bits/pixel), rather than using the full dynamic range of 65,536 grey levels (16 bits/pixel) before computing the MGLCM. This quantization step was essential to reduce the large number of zero-valued entries in the co-occurrence matrix (Kassner and Thornhill, 2010; Gomez et al., 2012). The computing time for implementing MGLCM for each slice was about 150 sec. by using an HP workstation Z820 with Xeon E5-3.8GHz (Quad-Core) and 16GB of RAM (random access memory).

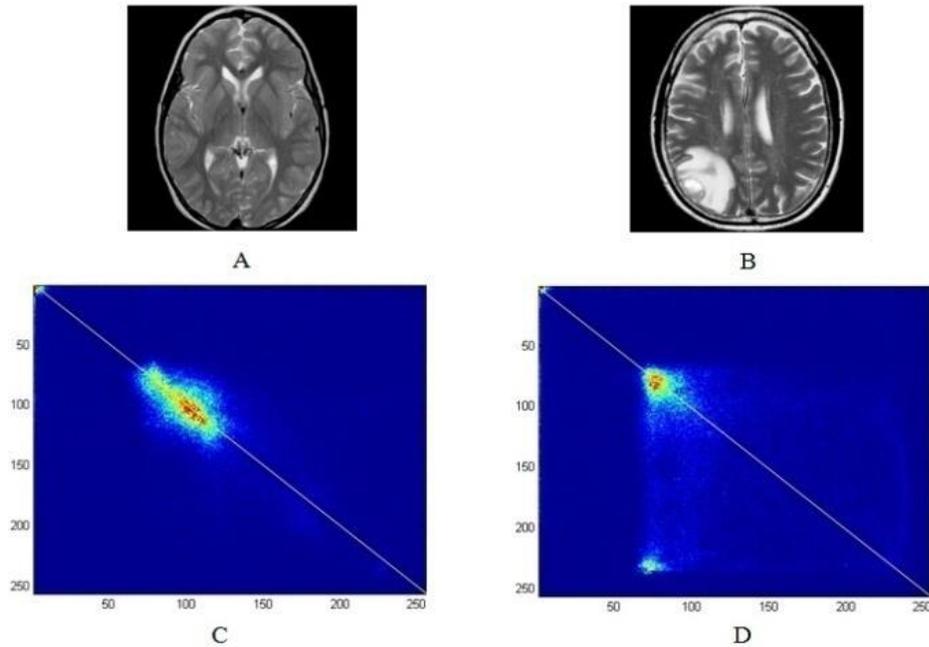


Figure 5.4: A) MRI normal brain slice, B) MRI abnormal brain slice, C) MGLCM of normal brain scanning and D) MGLCM of abnormal brain scanning.

Finally, nineteen texture descriptors are extracted from each co-occurrence matrix using the expressions given in the next section, representing the most common features derived from co-occurrence matrices (Haralick et al., 1973; Gomez et al., 2012; Birry, 2013; Yang et al., 2012; Hasan and Meziane, 2016; Bankman et al., 2009). These will be used in addition to the weighted mean and weighted distance predictors that are proposed in (Hasan and Meziane, 2016) and explained in detail in the next section. These predictors are used to measure statistically the degree of symmetry between the two hemispheres of the brain because the symmetry represents the main indicator in detecting pathological brains.

5.3 Texture Descriptions

In this study, nineteen texture descriptors that represent the most common descriptors are derived from co-occurrence matrices for each offset θ to classify a particular textures (Haralick et al., 1973; Gomez et al., 2012; Yang et al., 2012; Birry, 2013; Gadkari, 2004; Hasan and Meziane, 2016; Bankman et al., 2009). These are described in the following subsections in addition to the two newly proposed descriptors:

5.3.1 Contrast

The contrast descriptor is used to measure the local variations between the reference pixels in the left hemisphere of the brain and the opposite pixels in the right hemisphere of the brain. Where, the variation increases between both pixels when any abnormality or tumour appears in the brain and the increasing of variations lead to make the elements of the MGLCM are distributed away from the diagonal. It is defined by Eq. 5.2 (Haralick et al., 1973).

$$Contrast = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - j)^2 P(i, j) \quad 5.2$$

where P is the MGLCM matrix. i and j are the cell coordinates in the MGLCM and N is the number of grey levels used.

When i and j are equal, then the coefficient is located on the diagonal and according to the contrast equation there is not weight for all coefficients that are located on the diagonal. If i and j differ by 1, there is a very little displacement with weight equal to 1, and if they differ by 2, the weight is 4. Consequently, when there is a large amount of variation between the hemispheres of the brain the contrast will weight high value (Yang et al., 2012).

5.3.2 Correlation

The correlation descriptor is used to measure the strength of the linear dependencies of the reference pixels in the left hemisphere of the brain with opposite pixels in the right hemisphere. The correlation weights high values for normal brain scans and low values for pathological brain scans. It is defined by Eq. 5.3 (Yang et al., 2012; Haralick et al., 1973).

$$Correlation = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i, j) \frac{(i - \mu_x)(j - \mu_y)}{\sigma_x \sigma_y} \quad 5.3$$

where μ_x , μ_y , σ_x and σ_y are the means and the standard deviations of P_x and P_y respectively that are defined by Eq. 5.4, Eq. 5.5, Eq. 5.6, Eq. 5.7, Eq. 5.8 and Eq. 5.9:

$$P_x(j) = \sum_{i=0}^{N-1} P(i, j) \quad 5.4$$

$$P_y(i) = \sum_{j=0}^{N-1} P(i,j) \quad 5.5$$

$$\mu_x = \sum_{i=0}^{N-1} i \cdot P_x(i) \quad 5.6$$

$$\mu_y = \sum_{i=0}^{N-1} i \cdot P_y(i) \quad 5.7$$

$$\sigma_x = \sqrt{\sum_{i=0}^{N-1} (i - \mu_x)^2 \cdot P_x(i)} \quad 5.8$$

$$\sigma_y = \sqrt{\sum_{i=0}^{N-1} (i - \mu_y)^2 \cdot P_y(i)} \quad 5.9$$

5.3.3 Entropy

The entropy descriptor is used to measure the irregularity or complexity between the two hemispheres of the brain. It weights high values for pathological brain scans and low values for normal brain scans. It is defined by Eq. 5.10 (Pantelis, 2010).

$$Entropy = - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j) \log P(i,j) \quad 5.10$$

5.3.4 Energy

The energy predictor is used to measure the uniformity of MRI brain scans. It weights high values for normal brain scans and low values for normal brain scans. It is defined by Eq. 5.11 (Haralick et al., 1973; Yang et al., 2012).

$$Energy = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i, j)^2 \quad 5.11$$

5.3.5 Homogeneity

The homogeneity descriptor is used to measure the dissimilarity and contrast between both hemispheres of the brain. The dissimilarity leads to distribute the coefficients of MGLCM matrix away from the diagonal. The homogeneity descriptor is highly correlated with contrast descriptor. Such that, the homogeneity always weights values by the inverse of the contrast weight. It is defined by Eq. 5.12. (Yang et al., 2012; Haralick et al., 1973; Birry, 2013; Pantelis, 2010).

$$Homogeneity = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1 + (i - j)^2} \cdot P(i, j) \quad 5.12$$

5.3.6 Dissimilarity

The dissimilarity descriptor is used to measure the variation in the intensity value between the reference pixels in the left hemisphere and the opposite pixels in the right hemisphere. It weights high values for pathological brain scans and low values for normal brain scans. It is defined by Eq. 5.13 (Gomez et al., 2012; Gebejes and Huertas, 2013).

$$Dissimilarity = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |i - j| P(i, j) \quad 5.13$$

5.3.7 Sum of Square Variance

The sum of square variance descriptor weights high values for the coefficients which are differed significantly from the mean value of MGLCM. It is high for pathological brain scans and low for normal brain scans. It is defined by Eq. 5.14 (Gomez et al., 2012; Haralick et al., 1973; Pantelis, 2010).

$$\text{Sum of square variance} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - \mu)^2 \cdot P(i, j) \quad 5.14$$

5.3.8 Cluster Shade

The cluster shade descriptor is used to measure the skewness of the MGLCM matrix or the lack of symmetry. It is considered as a gauge of uniformity. It weights high values for pathological brain scans and low values for normal brain scans. It is defined by Eq. 5.15 (Yang et al., 2012; Haralick et al., 1973).

$$\text{Cluster shade} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i + j - \mu_x - \mu_y)^3 P(i, j) \quad 5.15$$

5.3.9 Cluster Prominence

The cluster prominence descriptor is also used to measure the skewness or asymmetry of the MGLCM. It weights high values for pathological brain scans and low values for normal brain scans. It is defined by Eq. 5.16 (Yang et al., 2012).

$$\text{Cluster prominence} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i + j - \mu_x - \mu_y)^4 P(i, j) \quad 5.16$$

5.3.10 Inverse Difference Normalized

The inverse difference normalized descriptor is used to measure the homogeneity of MRI brain scans. It weights relatively high values when the texture of the brain scans are homogeneous and small values when there is heterogeneity in the texture of the brain scans due to the abnormality. It is defined by Eq. 5.17 (Gomez et al., 2012).

$$\text{Inverse difference normalized} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1 + \frac{|i-j|}{N}} P(i, j) \quad 5.17$$

5.3.11 Inverse Difference Moment Normalized

The inverse difference moment normalized descriptor is also used to measure the homogeneity of MRI brain scan. It weights low values for pathological brain scans and high values for normal brain scans. It is defined by Eq. 5.18 (Gomez et al., 2012).

$$\text{Inverse difference moment normalized} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1 + \frac{(i-j)^2}{N}} P(i,j) \quad 5.18$$

5.3.12 Sum Average, Sum Entropy, Sum Variance and Difference Entropy

These descriptors weight high values for pathological brain scans and slightly low values for normal brain scans. The sum average, sum entropy, sum variance and difference entropy predictors are defined by Eq. 5.19, Eq. 5.20, Eq. 5.21 and Eq. 5.22 respectively (Pantelis, 2010; Gomez et al., 2012; Haralick et al., 1973; Albrechtsen, 2008).

$$\text{Sum average} = \sum_{i=1}^{2N-1} i P_{i+j}(i) \quad 5.19$$

$$\text{Sum entropy} = - \sum_{i=1}^{2N-1} P_{i+j}(i) \log P_{i+j}(i) \quad 5.20$$

$$\text{Sum variance} = \sum_{i=1}^{2N-1} (i - \text{sum entropy})^2 P_{i+j}(i) \quad 5.21$$

$$\text{Difference entropy} = - \sum_{i=1}^N P_{i-j}(i) \log P_{i-j}(i) \quad 5.22$$

where P_{i+j} and P_{i-j} are defined by Eq. 5.23 and Eq. 5.24 respectively.

$$P_{i+j}(k) = \sum_{i=1}^N \sum_{j=1}^N P(i,j) \quad , i + j = k \text{ and } k = 2,3, \dots \dots 2N - 1 \quad 5.23$$

$$P_{i-j}(k) = \sum_{i=1}^N \sum_{j=1}^N P(i, j) \quad , |i - j| = k \text{ and } k = 2, 3, \dots \dots N - 1 \quad 5.24$$

5.3.13 Information Measure of Correlation I and Correlation II

The information measure of correlation I descriptor is used to measure the linear dependency of grey levels of neighbouring pixels between both hemispheres of the brain. Furthermore, it is used to measure the deformation, displacement, strain and optical flow of a given MRI brain scan. It weights high values for pathological brain scans and slightly low values for normal brain scans. While the information measure of correlation II descriptor in the reverse state, it weights low values for pathological brain scans and high values for normal brain scans. They are defined by Eq. 5.25 and Eq. 5.26 (Gomez et al., 2012; Haralick et al., 1973).

$$\text{Information measure of correlation I} = \frac{HXY - HXY1}{\max(HX, HY)} \quad 5.25$$

$$\text{Information measure of correlation II} = \sqrt{1 - e^{(-2(HXY2 - HXY))}} \quad 5.26$$

where HX , HY , HXY , $HXY1$ and $HXY2$ are defined by Eq. 5.27, Eq. 5.28, Eq. 5.29, Eq. 5.30 and Eq. 5.31 respectively.

$$HX = - \sum_{i=0}^{N-1} P_x(i) \log P_x(i) \quad 5.27$$

$$HY = - \sum_{i=0}^{N-1} P_y(i) \log P_y(i) \quad 5.28$$

$$HXY = - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i, j) \log P(i, j) \quad 5.29$$

$$HXY1 = - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i, j) \log(P_x(i) P_y(j)) \quad 5.30$$

$$HXY2 = - \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_x(i)P_y(j) \log(P_x(i) P_y(j)) \quad 5.31$$

5.3.14 Autocorrelation

The autocorrelation descriptor is used to measure the linear spatial relationship between the hemispheres of the brain and estimate the intensity value concentration on MRI brain scanning image. It weights high values for pathological brain scans and relatively low values for normal brain scans. It is defined by Eq. 5.32 (Gomez et al., 2012).

$$Autocorrelation = \sum_{i=0}^N \sum_{j=0}^N (i . j)P(i, j) \quad 5.32$$

5.3.15 Maximum Probability

Maximum probability descriptor is the maximum value of the MGLCM coefficients. It is defined by Eq. 5.33 (Gomez et al., 2012).

$$Maximum\ probability = \max_{i,j} P(i, j) \quad 5.33$$

5.3.16 Weighted Mean

The weighted mean descriptor is proposed by Hasan and Meziane (2016) to detect the irregularity of MRI brain scans by calculating the nearest distance between the weighted mean to the diagonal of the MGLCM as shown in Fig. 5.5. It weights high values for pathological brain scans and low values for normal brain scans. It is defined by Eq. 5.34 and Eq. 5.35 (Hasan and Meziane, 2016).

$$x = \frac{1}{256^2} \sum_{i=1}^{256} \sum_{j=1}^{256} i . P(i, j) \quad 5.34$$

$$y = \frac{1}{256^2} \sum_{i=1}^{256} \sum_{j=1}^{256} j . P(i, j) \quad 5.35$$

where x and y are the coordinates of weighted mean in the MGLCM as given in Eq. 5.36.

$$\text{Weighted mean} = |y - x| \sin 45^\circ \quad 5.36$$

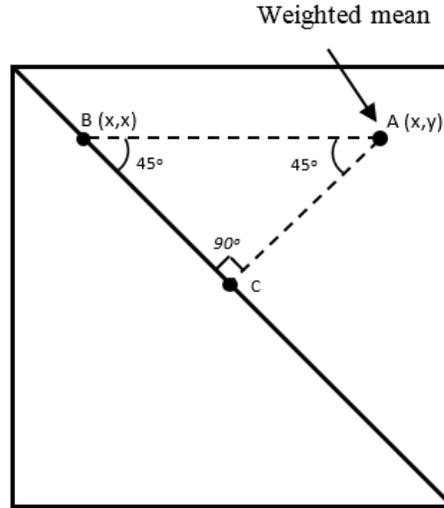


Figure 5.5: Weighted mean representation on MGLCM.

5.3.17 Weighted Distance

Weighted distance descriptor is also proposed by Hasan and Meziane (2016) to detect the irregularity of MRI brain scan by multiplying each coefficient in the MGLCM by the nearest distance d to the diagonal as shown in Fig. 5.6. It weights high values for pathological brain scans and low values for normal brain scans. It is defined by Eq. 5.37 and Eq. 5.38 (Hasan and Meziane, 2016).

$$\text{upper}_{\text{triangular}} = \sum_i \sum_j d_{ij} \cdot P(i, j) \quad 5.37$$

where i and j are the elements' coordinates that locate in the upper triangular of MGLCM.

$$\text{lower}_{\text{triangular}} = \sum_i \sum_j d_{ij} \cdot P(i, j) \quad 5.38$$

where i and j are the elements' coordinates that locate in the lower triangular of MGLCM.

Then the weighted distance is determined by Eq. 5.39.

$$\text{Weighted distance} = |\text{upper}_{\text{triangular}} - \text{lower}_{\text{triangular}}| \quad 5.39$$

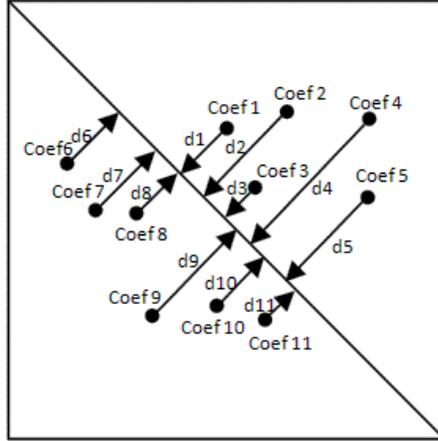


Figure 5.6: Weighted distance representation on MGLCM.

5.3.18 Cross-Correlation Coefficients

Cross-Correlation Coefficient represents one of the most important and most useful statistics for measuring the similarity between two sets of data. Usually, it is quantified by a single number that is called the correlation-coefficient (r), between brain hemispheres. Such that the left hemisphere of brain L is the reference image and the right hemisphere of brain R is the other image. It is defined by Eq. 5.40 (Birry, 2013).

$$r = \frac{1}{(n-1)} \sum_n \left(\frac{(L_n - \bar{L})}{\sigma_L} \times \frac{(R_n - \bar{R})}{\sigma_R} \right) \quad 5.40$$

where n is the total number of pixels in the brain hemisphere image. \bar{L} and \bar{R} are the means of left and right hemisphere images respectively. σ_L and σ_R are the standard deviations of the left and right hemisphere images respectively. For strong similarity between both hemispheres of the brain, the correlation coefficient approximates to 1 and decreases significantly to 0 when there is abnormality that appears in one or both hemispheres of the brain (Birry, 2013).

5.4 Feature Aggregation

The MGLCM method determines nine co-occurrence matrices. For each matrix, 21 statistical descriptors are determined, generating 189 descriptors for each MRI brain scan (Hasan and Meziane, 2016). The cross-correlation descriptor is also determined for the original MRI brain scan. Accordingly, 190 descriptors are attained for each MRI brain scan. These features are used by the subsequent classification to differentiate between normal and abnormal MRI brain scans.

5.5 Feature Selection

High dimensional feature sets can negatively affect the classification results because high number of features may reduce the classification accuracy owing to the redundancy or irrelevance of some features (Babatunde et al., 2014). Feature-selection techniques aim to identify a small subset of features that minimizes redundancy and maximizes relevancy (Tang et al., 2014). Therefore, feature selection is an important step in exposing the most informative features and for optimally tuning the classifier's performance to reliably classify unknown data (Pantelis, 2010).

In this study, ANOVA was employed to measure feature significance and relevance as explained in details in section 2.4.1.1. It is a robust statistical technique used for data analysis and for detecting the level of significance of each predictor in the feature. The critical value α is set at 0.001 to obtain highly significant features (Johnson and Synovec, 2002). The assessment of predictors depends on both F-statistic value and P-value because a P-value less than 0.001 is insufficient for measuring significance of a predictor. Instead, the predictor must also hold a high F-statistic value. The high F-statistic value indicates that the classes are significantly separated from one another (Hasan and Meziane, 2016). The differences between the features of normal and abnormal MRI brain scan groups of the co-occurrence matrix at $\theta_1=0$ and $\theta_2=0$ is shown in Table 5.1. All features seemed acceptable except the weighted mean descriptor. Nevertheless, significant variation existed in the F-statistic values between features, indicating a degree of significant difference between the selected features.

The P-value does not actually signify the degree of separation of each group from others and ignores feature redundancy (Johnson and Synovec, 2002). This drawback is overcome by using the F-statistic to determine the power of feature discrimination through thresholding. Such that, different threshold values are taken to ignore the redundant features and evaluate the selected features at each time by observing the performance of the classifier. When the F-statistic threshold value increases, the numbers of selected descriptors and the vector of the features decrease. The ANOVA was implemented by using IBM SPSS Statistics software Version 20 (Burns and Burns, 2008).

5.6 Feature Normalization

It is noted, that there is a variation in the ranges from the extracted descriptors. These differences in ranges can lead to making some descriptors that have large values influence more than other descriptors with small values on the behaviour of the classifier (Larose, 2005). Therefore, data normalization is an essential step that prepares and normalizes the given descriptors to standardize the scale of effect of each descriptor. It helps to improve the performance of the classifier by transforming the given raw descriptors into better form and more suitable for the training process (Jayalakshmi and Santhakumaran, 2011). In this study, min-max normalization approach is used to perform a linear transformation on the extracted descriptors with preserving the relationships between the original descriptors. It was described in details in section 2.4.2.1.

5.7 Feature Classification

For classification, the three most common supervised classification techniques are used in this study to classify the extracted features. These classifiers are LDA (Zacharaki et al., 2009; Takayanagi et al., 2011), SVM (Nagarajan et al., 2013; Kharrat et al., 2010; Bauer et al., 2011; Hackmack et al., 2012; Kalbkhani et al., 2013) and MLP (Pantelis, 2010; Saritha et al., 2013; Antkowiak, 2006). The confusion matrix is a useful tool used for analysing the performance of the classifiers. It is a matrix of size m by m , that allows to assess and describe the classification performance, where m denotes number of classes (Han et al., 2011). In this study, there are two classes; the positive class refers to the pathological brain scans and the negative class refers to the normal brain scans.

Table 5.1: Comparison of MRI brain scans features (mean \pm standard deviation (SD)) between normal and abnormal patients

Features	Abnormal MRI scans	Normal MRI scans	F-statistic	P-value
Auto correlation ($\times 10^3$)	5.626 \pm 1.2	4.92 \pm 1.24	13.67	<0.001
Contrast ($\times 10^3$)	1.89 \pm 0.618	0.918 \pm 0.229	166.2	<0.001
Correlation ($\div 10$)	7.1 \pm 0.91	8.07 \pm 0.72	291.5	<0.001
Cluster Prominence ($\times 10^8$)	3.64 \pm 1.87	2.7 \pm 1.09	14.62	<0.001
Cluster Shade ($\times 10^5$)	7.6 \pm 4.26	5.5 \pm 2.9	13.14	<0.001
Dissimilarity ($\times 10$)	2.42 \pm 0.47	1.58 \pm 0.21	209	<0.001
Energy ($\div 10$)	1.022 \pm 0.2	1.05 \pm 0.18	368.15	<0.001
Entropy	7.07 \pm 0.336	6.87 \pm 0.25	15.21	<0.001
Homogeneity ($\div 10$)	3.55 \pm 0.34	3.76 \pm 0.26	451.3	<0.001
Max. Probability ($\div 10$)	3.178 \pm 0.33	3.23 \pm 0.28	444.96	<0.001
Sum of Square Variance ($\times 10^3$)	6.5 \pm 1.6	5.384 \pm 1.23	24.36	<0.001
Sum Average ($\times 10^2$)	1.15 \pm 0.112	1.06 \pm 0.15	20.84	<0.001
Sum Variance ($\times 10^4$)	2.337 \pm 0.47	1.97 \pm 0.48	24.25	<0.001
Sum Entropy	4.46 \pm 0.177	4.16 \pm 0.147	35.98	<0.001
Difference Entropy	3.64 \pm 0.2	3.34 \pm 0.124	132.2	<0.001
Information Measure of Correlation I ($\div 10$)	-2.24 \pm 0.3	-2.53 \pm 0.26	430.15	<0.001
Information Measure of Correlation II ($\div 10$)	9.11 \pm 0.2	9.26 \pm 0.18	355.48	<0.001
Inverse Difference Normalized ($\div 10$)	9.25 \pm 0.12	9.48 \pm 0.06	407.8	<0.001
Inverse Difference Moment Normalized ($\div 10$)	9.78 \pm 0.07	9.87 \pm 0.028	316.89	<0.001
Weighted Mean ($\div 10$)	-8.73 \pm 84	0.53 \pm 18.7	0.92	0.339
Weighted Distance	3.05 \pm 2.91	0.77 \pm 0.52	46.1	<0.001
Cross Correlation ($\div 10$)	7.1 \pm 0.91	8.07 \pm 0.72	291.5	<0.001

So, the confusion matrix includes the following four terms; TP (true positive), FP (false positive), TN (true negative) and FN (false negative). Such that, TP is the patients who are correctly classified as pathological, TN is the patients who are correctly classified as healthy, FN is the patients who are incorrectly classify as pathological and FP is the patients who are incorrectly classified as healthy. Good performance corresponds to large numbers down the main diagonal of confusion matrix and ideally tend to be zero on its off diagonal (Witten et al., 2011).

Set of statistical measures that are useful in analysing, evaluating and quantifying a classifier's performance: the overall accuracy is given in Eq. 5.41 (Witten et al., 2011).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad 5.41$$

The accuracy is not sufficient to evaluate the performance of a classifier, the sensitivity and specificity measures can be used respectively to assess how the classifier can discriminate pathological patients (positive class) and how it can discriminate normal patients (negative class). Sensitivity refers to the TP rate that is proportional to the pathological patients that are correctly classified, while the specificity refers to the TN rate that is proportional to the healthy patients that are correctly classified (Han et al., 2011), these measures are calculated using Eq. 5.42, and Eq. 5.43.

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad 5.42$$

$$Specificity = \frac{TN}{TN + FP} \times 100\% \quad 5.43$$

In this study, the collected dataset is randomly partitioned into k folds D_1, D_2, \dots, D_k , that are approximately of equal size. The training and testing are repeated k times and the classification result is determined as the average of the overall classification accuracies (Dubitzky et al., 2007). For instance, in the first iteration, partition D_1 is reserved for testing and the remaining partitions D_2, D_3, \dots, D_k are reserved collectively for training the model and so on (Han et al., 2011). Each fold is held out only one time for testing and $k-1$ for learning. In general, 10-fold and 5-fold cross validation are recommended for estimating the accuracy due to its relatively low bias and low variance (Tantisatirapong, 2015; Han et al., 2011; Nabizadeh and Kubat, 2015; Birry, 2013). The cross-validation was implemented by using *crossvalind* function with *kfold* in MATLAB R2013a Image Processing Toolkit (Matlab, 2013). Figure 5.7 shows the flowchart and steps of training and testing of the classification techniques.

As mentioned in chapter four, the collected dataset in this study includes MRI brain scans of 165 cases and it was clinically classified into normal and abnormal by the clinicians. The highest classification accuracy with the best performance was achieved using the MLP network at 91% while LDA and SVM achieved 77% and 87% respectively. The LDA is implemented using *classify* function and the SVM is implemented using *svmtrain* and *svmclassify* functions in MATLAB R2013a Image Processing Toolkit (Matlab, 2013). While the MLP requires to be configured optimally by setting the number of hidden layers and the number of neurons in the hidden layer as explained in details in sections 2.5.3.2.

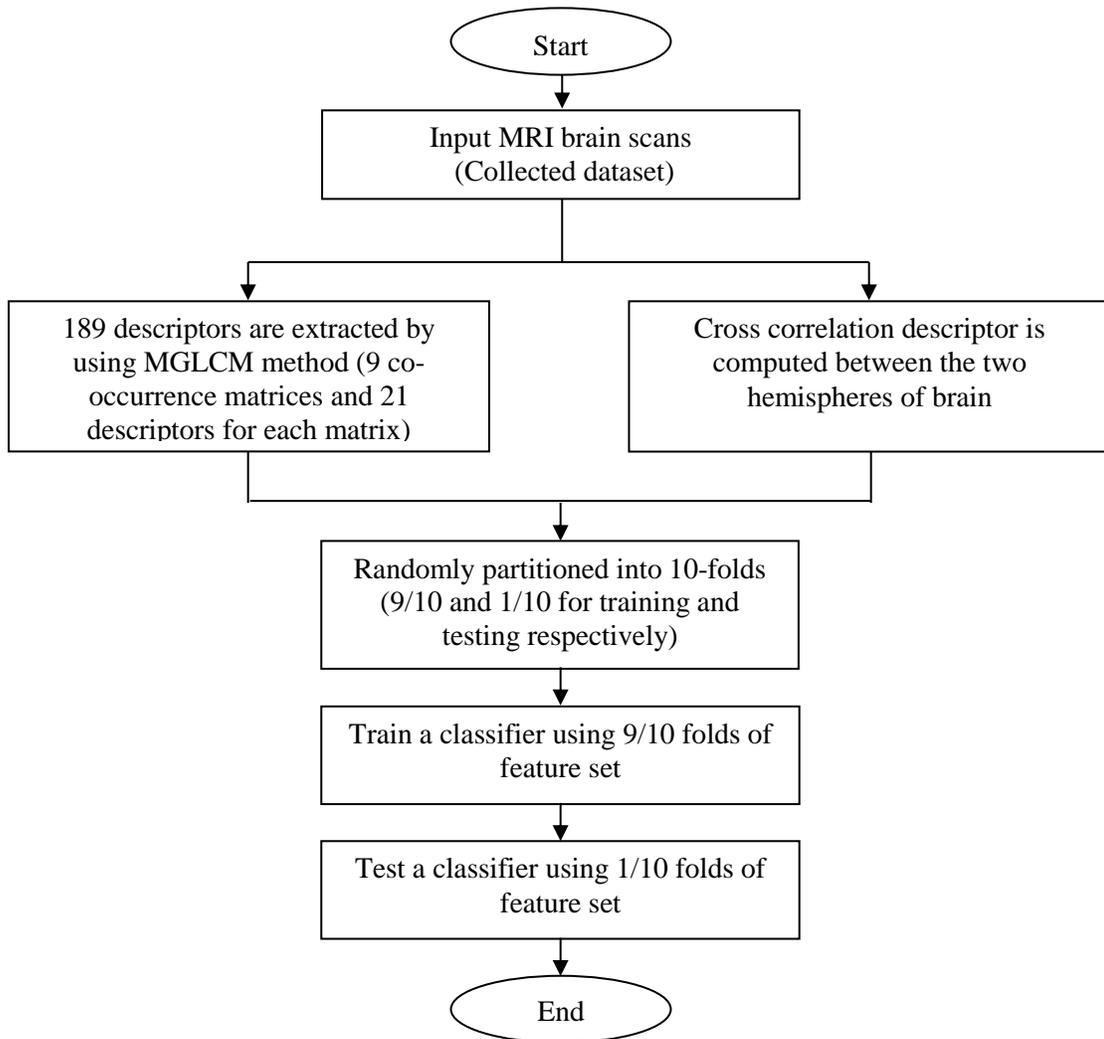


Figure 5.7: Flowchart and steps of training and testing of classification techniques.

The number of neurons in the input layer is usually equal to the number of descriptors in the input feature vector; therefore, it is set to 190. The output layer of MLP network is set to one neuron because we have only two classes (normal and abnormal brain scans). The number of hidden layers is chosen approximately by the user and generally, one hidden layer is sufficient for approximating and classifying MRI brain scans. While the number of neurons in the hidden layer is tuned by changing the number of neurons in the hidden layer every experiment until getting the lowest root mean square error (RMSE) value (Gunasekara et al., 2009), as defined using Eq. 5.44.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (g_i - y_i)^2}{n}} \quad 5.44$$

where \mathbf{g} denotes the actual predicted value by neural network, \mathbf{y} denotes the target value of the given dataset and \mathbf{n} is the number of observation in the dataset. The *newff* function in (Matlab, 2013) is used to create and train the MLP network. The transfer function that is used is the sigmoid function and the training function that is used to update weights and bias values is the scaled conjugate gradient method (*trainscg*). It is faster than the default function (*trainlm*) for larger datasets. The target values are set to 0 and 1, where 0 denotes the normal brain scans and 1 denotes the abnormal brain scans. In order to decide the number of neurons in the hidden layer, 10 runs with different number of neurons in hidden layer are implemented and the RMSEs of runs are shown in Table 5.2. Figure 5.8 shows the mean values of RMSE for runs and it can be clearly seen that the minimum mean value occurred at 75 neurons in hidden layer with average value of RMSE is 0.2971. Therefore, in this study the number of neurons in hidden layer is set at 75 neurons to achieve the maximum classification accuracy by MLP network and the training of MLP network with this configuration is shown in Fig. 5.9.

The same co-occurrence statistics which were used in this study were computed using the traditional GLCM method in order to perform a comparison between the two methods. Four co-occurrence matrices with four orientations (0° , 45° , 90° and 135°) and distance 1 were computed. The maximum classification accuracy was achieved by MLP at 86%, followed by 82% for SVM and the LDA classifier achieved 74%. Additionally, the performance of the features which were extracted by MGLCM was compared with the Gabor wavelet that was applied with five different scales and eight orientations.

Table 5.2: The results of 10 runs for 190 input descriptors with different number of neurons in hidden layer.

Neurons No.	RMSE of Runs										Average
	1	2	3	4	5	6	7	8	9	10	
50	0.35	0.36	0.43	0.335	0.267	0.3	0.285	0.47	0.32	0.3	0.3417
75	0.285	0.226	0.285	0.335	0.32	0.335	0.3	0.3	0.285	0.3	0.2971
100	0.378	0.226	0.35	0.36	0.267	0.335	0.247	0.285	0.3	0.285	0.3033
125	0.463	0.335	0.267	0.285	0.428	0.335	0.3	0.378	0.335	0.3	0.3426
150	0.226	0.226	0.364	0.416	0.378	0.416	0.378	0.35	0.247	0.3	0.3301

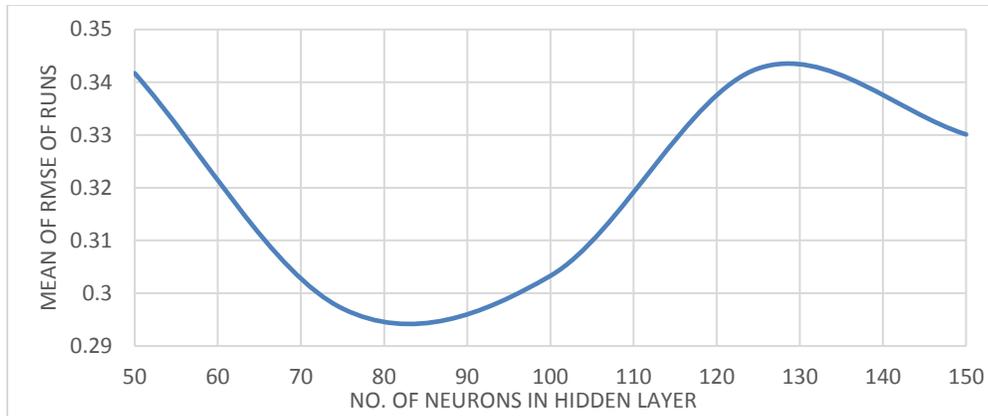
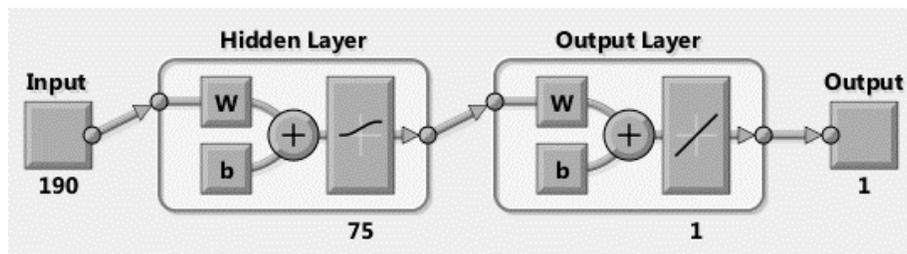
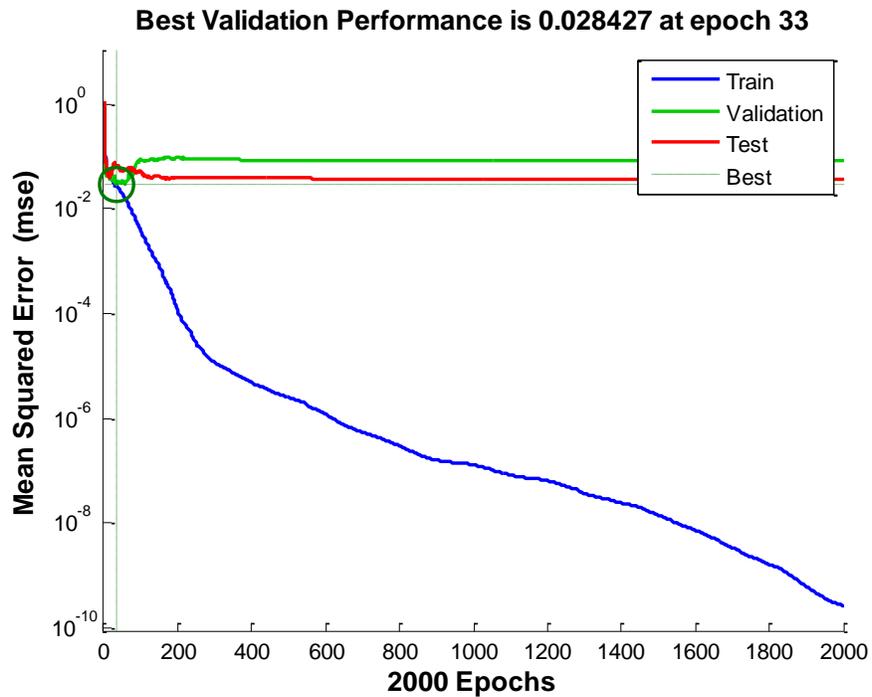


Figure 5.8: Number of neurons vs. the mean of RMSE of runs.



A



B

Figure 5.9: A) MLP network structure and B) The performance of MLP network.

The length of the Gabor feature vector was 655360 features for each MRI brain scans. The achieved classification accuracy by the three classifiers was 90% by SVM, 87.4% by MLP and 62.5% by LDA (Hasan et al., 2016c). It is noted that there is a prevalence in classification accuracy of texture features that were extracted using MGLCM than those extracted using GLCM and Gabor wavelet in both LDA and MLP classifiers. While in SVM, the performance of the extracted features using Gabor wavelet outperforms those achieved by others as shown in Fig. 5.10.

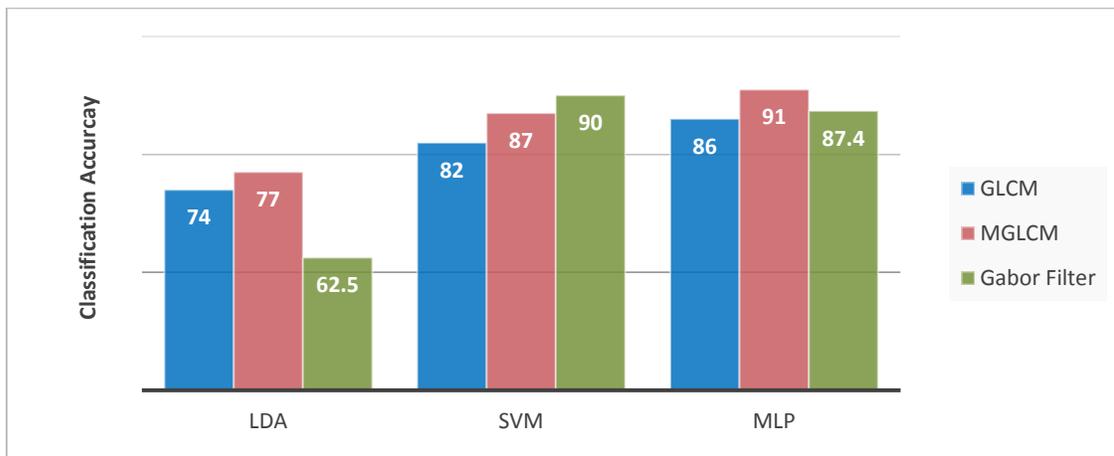


Figure 5.10: A comparison between MGLCM, GLCM and Gabor wavelet regards classification accuracy.

It is beneficial to select the most significant extracted features by MGLCM in order to improve the classifier performance by reducing misclassified data. ANOVA method was implemented with different F-statistic threshold values as a feature selection method. At each run the descriptors which have an F-statistic value greater than the threshold value, were determined. The results of the classification are shown in Table 5.3.

Table 5.3: The result of using four different F-statistic threshold values on the classification accuracy.

F-statistic threshold value	No. of selected descriptors	LDA Accuracy	SVM Accuracy	MLP	
				Neurons in hidden layer	Accuracy
0	181	80%	87%	85	92.5%
30	127	81%	95%	60	96.6%
35	100	83%	96%	55	97.8%
50	55	82%	95%	25	94%

It is noted that the best performance of the three classifiers have improved significantly when the threshold value is experimentally fixed to 35, where the number of descriptors were reduced to 100 as shown in Fig. 5.11. The achieved classification accuracies of LDA and SVM classifiers have increased to 83% and 96% respectively. While, the average performance of the MLP network was $97.8 \pm 0.1\%$. The sensitivity and specificity rates were $98.1 \pm 0.3\%$ and $97.6 \pm 0.4\%$ respectively.

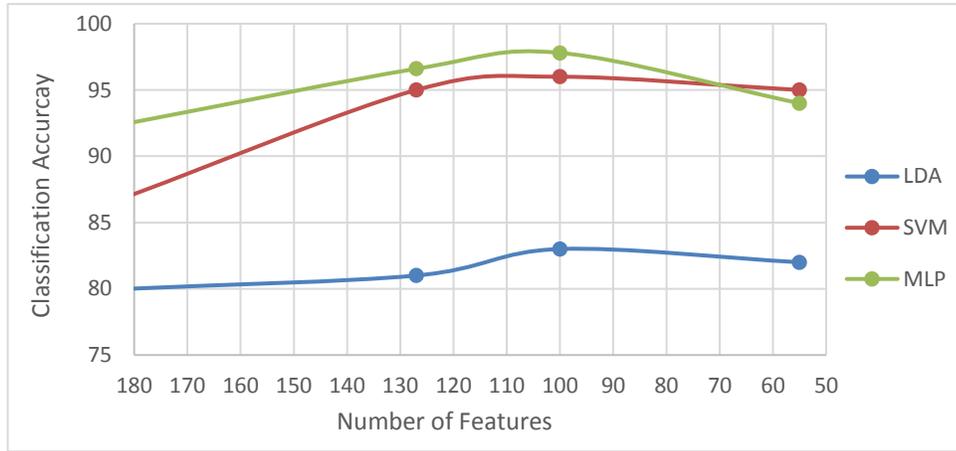


Figure 5.11: The optimal number of features corresponding with classification accuracy.

The transfer function used in the MLP network was the tangent function. To update the weights and bias value, the scaled conjugate gradient method (*trainscg*) in MATLAB R2013a Image Processing Toolkit (Matlab, 2013), was used to be faster than the default function (*trainlm*) for larger datasets. In order to decide the number of neurons in the hidden layer, 10 runs with different number of neurons in hidden layer were implemented and the RMSEs of runs are shown in Table 5.4.

Table 5.4: The results of 10 runs for 100 input descriptors with different number of neurons in hidden layer.

Neurons No.	RMSE of Runs										Average
	1	2	3	4	5	6	7	8	9	10	
40	0.322	0.288	0.3	0.35	0.35	0.25	0.25	0.322	0.202	0.27	0.29
45	0.456	0.247	0.32	0.32	0.285	0.322	0.25	0.288	0.322	0	0.28
50	0.322	0.204	0.247	0.204	0.202	0.144	0.202	0.143	0.35	0.204	0.22
55	0.13	0.13	0.1	0.12	0.11	0.144	0.144	0.13	0.14	0.143	0.129
60	0.25	0.202	0	0.288	0.144	0.143	0.202	0.204	0.202	0.143	0.178
65	0.19	0.27	0.288	0.23	0.16	0.12	0.2	0.25	0.202	0.23	0.214

Figure 5.12 shows the average values of RMSE for runs and it can be clearly seen that the minimum value of RMSE occurred at 55 neurons in hidden layer with average value of 0.129. The training of the MLP network with 100 input layer neurons, 55 hidden layer neurons and single output layer is shown in Fig. 5.13.

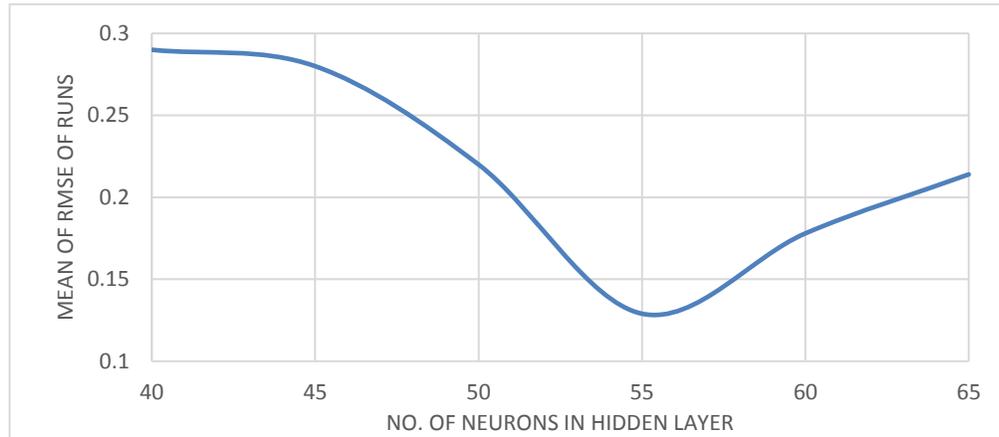
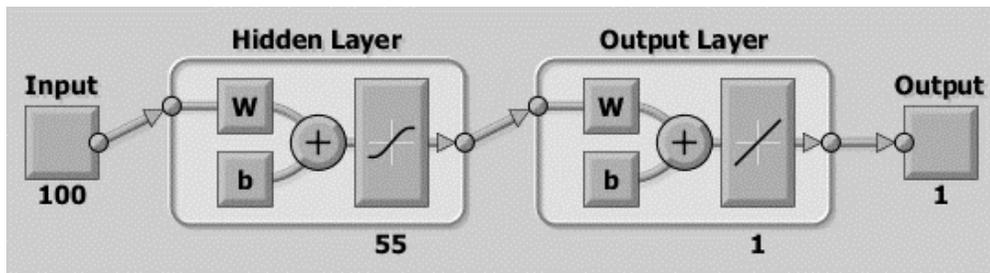


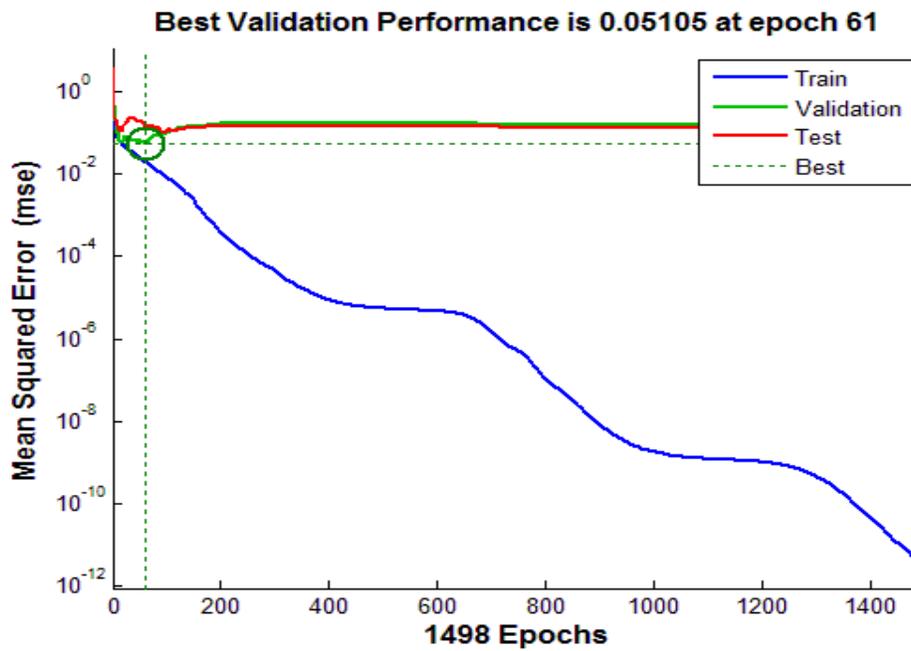
Figure 5.12: Number of neurons vs. RMSEs of runs.

Consequently, the number of descriptors in the feature vector is reduced from 190 to 100. Where, 90 descriptors were discarded and considered as irrelevant or redundant features, and 11 relevant and significant descriptors for each angle of the MGLCM are selected by ANOVA, namely, *contrast*, *correlation*, *dissimilarity*, *sum of square variance*, *sum average*, *sum variance*, *difference entropy*, *information measure of correlation I*, *inverse difference normalized (IDN)*, *inverse difference moment normalized (IDMN)* and *weighted distance* in addition to the *cross correlation*. Figure 5.14 shows how these selected descriptors are significantly different in means and standard errors (SE) between the two groups.

The ANOVA method was also applied on the texture features which were extracted by the traditional GLCM. The best classification accuracy was achieved when the F-statistic threshold value was 35, such that the selected texture features were *auto correlation*, *cluster prominence*, *cluster shade*, *sum of square variance*, *sum variance* and *cross correlation*. The maximum classification accuracy with the best performance of GLCM achieved was 92% for MLP, 90% for SVM and 79% for LDA. While, the maximum classification accuracy with the best performance of Gabor features achieved was 92% for MLP, 94% for SVM and 79% for LDA (Hasan et al., 2016c).



A



B

Figure 5.13: A) MLP network structure and B) The performance of MLP network.

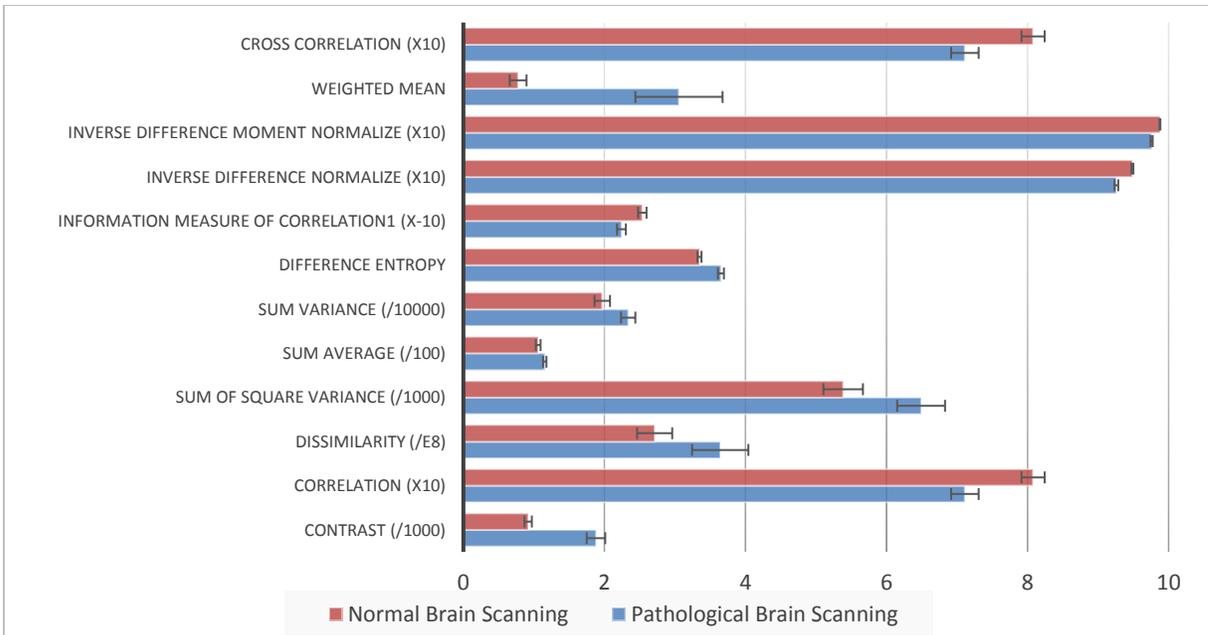


Figure 5.14: Textural features (mean \pm standard error (SE)) of the normal and pathological MRI brain scans.

It can be clearly seen that there is a prevalence in classification accuracy of texture features that were extracted using MGLCM than those extracted using GLCM and Gabor wavelet in all classifiers as shown in Fig. 5.15.

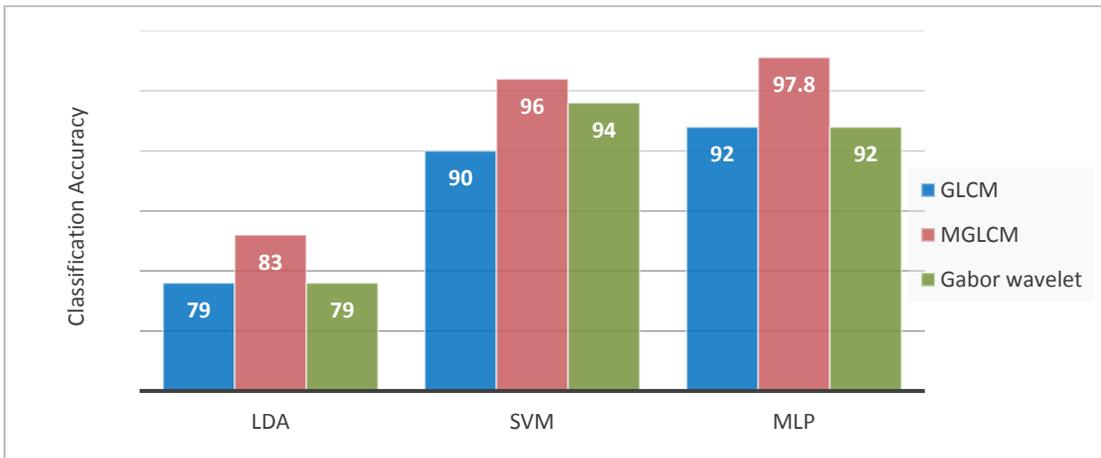


Figure 5.15: A comparison of classification accuracies for MGLCM, GLCM and Gabor wavelet using ANOVA method.

The performance of ANOVA is validated by comparing its results against those obtained by other techniques for feature selection mainly PCA and Kernel PCA methods which were implemented and used by (Van der Maaten, 2007; Moghaddasi et al., 2014). Table 5.5 and Table 5.6 show the results of classification when using PCA and Kernel PCA respectively with same dimensions that was used in ANOVA (127, 100 and 55).

Table 5.5: Classification accuracy for the three classifiers with PCA method for feature selection.

No. of selected predictors	LDA Accuracy	SVM Accuracy	MLP	
			Neurons in hidden layer	Accuracy
127	65%	78%	65	66%
100	63%	80%	55	77%
55	63%	78%	25	78%

Table 5.6: Classification accuracy for the three classifiers with Kernel PCA method for feature selection.

No. of selected predictors	LDA Accuracy	SVM Accuracy	MLP	
			Neurons in hidden layer	Accuracy
127	60%	81%	65	62%
100	61%	74.5%	55	67%
55	58%	76%	25	72%

The Correlation Feature Selection Subset Evaluator (CFS-SE) method was also used for the evaluation and validation of the extracted features. This method was implemented using the WEKA software (Witten et al., 2011; M. Hall, 1999). In total, six descriptors for all angles were chosen as the most relevant by this method and they are *weighted distance* (0,0), *sum of square variance* (0,315), *dissimilarity* (315,45), *inverse difference normalize* (315,45), *inverse difference normalize* (315,315) and *inverse difference moment normalize* (45,315). Table 5.7 summarizes the result of the classification of the selected features using the WEKA CFS-SE method with the same classifiers that were used previously.

Table 5.7: Classification accuracy for the three classifiers with WEKA CFS-SE method for feature selection.

No. of selected predictors	LDA Accuracy	SVM Accuracy	MLP	
			Neurons in hidden layer	Accuracy
6	75%	89%	3	91%

Figure 5.16 summarizes the behaviour of the implemented features selection methods and shows that the ANOVA outweighs the performance of the other techniques for detecting the most relevant predictors.



Figure 5.16: Performance of the implemented feature selection methods.

The experimental results of the proposed algorithm are compared with previous studies as shown in Table 5.8 (Hasan and Meziane, 2016).

Table 5.8: Comparison with previous proposed methods

Reference	Features methods	No. of Patients	Classifier	Accuracy
(Nabizadeh and Kubat, 2015)	- First-order statistical - GLCM (4 orientations and 2 distances) - GLRLM (4 orientations) - HOG - LBP	25 (BRATS 2013)	SVM	97.4%
(Gomez et al., 2012)	- GLCM (4 orientations and 10 distances)	436	LDA	87%
(Sachdeva et al., 2016)	- GLCM (4 orientations and 1 distances) - LoG - DGTF - RICGF - RILBP - IBF - SBF	55	SVM	91.7%
			MLP	94.9%
Proposed algorithm	- MGLCM (9 orientations and 1 distance)	165	MLP	97.8%
		25 (BRATS 2013)		98.6%

Nabizadeh and Kubat (2015) used five methods for feature extraction; first-order statistical features, GLCM with four orientations (0° , 45° , 90° and 135°) and two distances (1 and 2), GLRLM with four orientations (0° , 45° , 90° and 135°), HOG and linear LBP. The feature vector included 475 descriptors. The dataset included 25 patients (BRATS 2013). A classification accuracy of 97.4% was achieved by SVM.

Gomez et al. (2012) used the GLCM method for feature extraction to classify breast ultrasound images (BUS) with 22 descriptors computed from four co-occurrence matrices with orientations (0° , 45° , 90° and 135°), ten distances (1-10 pixels) and six quantization levels (8, 16, 32, 64, 128 and 256). To reduce the dimensionality of the feature space, the texture descriptors of the same distance were averaged over all orientations from 880 to 220. Additionally, mutual information (MI) is used for evaluating the quality of the features subset. The maximum classification accuracy was achieved by the LDA classifier at 87% for classifying 436 BUS images. The selected descriptors were feature / θ /d; correlation I/90/8, cluster prominence /0/1, correlation II/90/8, contrast /90/1, correlation I/90/9, difference variance /90/1, correlation II/90/9, correlation I/90/2, correlation I/90/7, inverse difference moment normalize/90/1, correlation II/90/7, correlation I/90/10, correlation I/90/6, correlation II/90/10, correlation II/90/6, correlation I/90/5 and inverse difference moment normalize /90/1.

The automated screening system in this study (Hasan and Meziane, 2016) depends essentially on the single proposed method for texture feature extraction MGLCM with nine orientations. The significant features were selected using ANOVA method and reduced to 100 descriptors. Over the entire collected dataset which included 165 patients and the standard dataset (BRATS 2013), the average achievable accuracy was 97.8% and 98.6% respectively by using MLP.

5.8 Conclusion

Since the visual diagnosis of the MRI scans is subjective and depends on the expertise of the radiologist, texture analysis has been widely studied for improving the diagnosis of MRI brain scans. In this study, 19 co-occurrence statistics which were most popular and common in previous studies, in addition to two proposed descriptors (weighted distance and weighted mean), were extracted from nine MGLCM matrices to discriminate brain abnormalities. Only 11 co-occurrence statistics, in addition to cross correlation descriptor, were chosen as the most significant features by ANOVA. The weighted distance feature was included within these 11

co-occurrence statistics and was chosen by ANOVA as a significant feature. ANOVA contributed to improve the behaviours of the classifiers in this study, the highest classification accuracy was $97.8 \pm 0.1\%$, and was achieved by combining MLP network with ANOVA method by taking only 100 relevant descriptors with 55 neurons in the hidden layer. Compared to the highest classification accuracy of combining SVM with PCA method which was 80% and 81% by combining the SVM with Kernel PCA method. Finally, 91% was achieved by combining the MLP with the WEKA CFS-SE method.

It is concluded, that the statistical texture features which were extracted by MGLCM are sufficient to discriminate the pathological patients from non-pathological patients by using T2-w images because most of the brain tumours appear hyper-intense in these images relative to normal brain tissue. A further advantage of our approach is that it uses a single MRI scan modality (T2-w image). The MGLCM gives high performance and accuracy in discriminating the normality and abnormality of the brain. However, the method is computational expensive and memory requirements represent the main disadvantages. Therefore, we will try to find a new method with less computation time and it will be demonstrated in next chapter.

CHAPTER SIX

Three-Dimensional Modified Grey Level Co-occurrence Matrix

Overview

This chapter covers the implementation of the proposed 3DMGLCM method for texture feature extraction in addition to comparing different techniques for classification such as LDA, SVM and ANN. Finally, a comparison between 2DMGLCM and 3DMGLCM was concluded.

6.1 Introduction

Medical image analysis is a rewarding field for investigating, developing and applying methods of image processing, computer vision and pattern recognition. Medical images are different from other images, as they show distributions of various physical features measured from the human body and attributes that are not otherwise accessible (Toennies, 2012). Many medical technologies show a projection of the three-dimensional human body onto a two-dimensional (2D) plane and slice images in such a way that the slices may be stacked to create a volume model. Most of the clinicians' time is spent on data examination and interpreting medical images. Therefore, it is essential to have a high level of experience to carry out manual and accurate segmentation and classification of these images in order to achieve the final diagnosis. However, due to the large number of slices which are produced by medical scanners, the manual detection of tumours is considered to be very cumbersome, a time consuming task and prone to human errors (Mortazavi et al., 2012; Menze et al., 2015). Texture analysis is one of the major features extraction techniques to identify RoI in an image or to classify an object (Rahim et al., 2014). Most of the proposed anatomical feature extraction techniques utilize 2D texture analysis, and the 2D approaches have some difficulties especially when the major axis of the desired object is not perpendicular to the image plane or there might be a degree of skewness (Ashton et al., 1997). Although the 2D approaches are relatively fast, they might lose relevant information present in neighbouring slices of MRI, contribute to obtain high dimensional feature vector because they aggregate feature from multiple slices are used for implementing 2D approaches. In this study, the concentration will be on textural feature in a three-dimensional

scheme from all MRI slices at the same time in order to reduce the computation time and complexity. The texture features will be extracted from MRI brain slices to encode clinically valuable information by the proposed 3DMGLCM method that will be used to measure statistically the similarity between the two separated volumetric hemispheres of brain.

Initially, prior pre-processing algorithms that should be used to prepare the MRI brain slices for texture features extraction by 3DMGLCM are explained in details in section 5.2 and illustrated in Fig. 6.1.

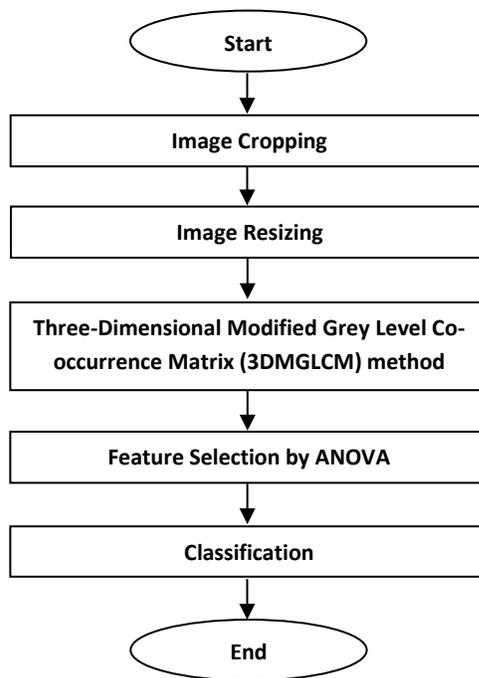


Figure 6.1: Flowchart of the implementation of the three-dimensional feature extraction, selection and classification.

6.2 Three-Dimensional Modified Grey Level Co-occurrence Matrix

Three-dimensional modified grey level co-occurrence matrix (3DMGLCM) method gives information about the patterning of the texture of MRI brain scans which could be used to calculate textural features. These features are extracted from volumetric data of MRI brain scan and used to measure statistically the degree of symmetry between the two hemispheres of the brain. It is a second order statistical method, used to generate texture features of MRI brain

scans by computing a spatial relationship of the joint frequencies of all pairwise combinations of grey levels configuration of each pixel in the left hemisphere of brain, which is considered as a reference pixel, with one of the nine opposite pixels in the right hemisphere according to the nine offsets. These nine pixels are distributed among three opposite successive slices according to the nine offsets $\theta = (45,45), (0,45), (315,45), (45,0), (0,0), (315,0), (45,315), (0,315), (315,315)$, and one distance $d=1$. Figure 6.2 shows how the joint frequencies of all pairwise combinations of grey levels configuration of reference pixel in slice z with nine opposite pixels which are distributed over slices $z+1, z$ and $z-1$. Consequently, because each pixel on the left hemisphere has nine opposite pixels on the right hemisphere, nine co-occurrence matrices are determined for each MRI brain scan.

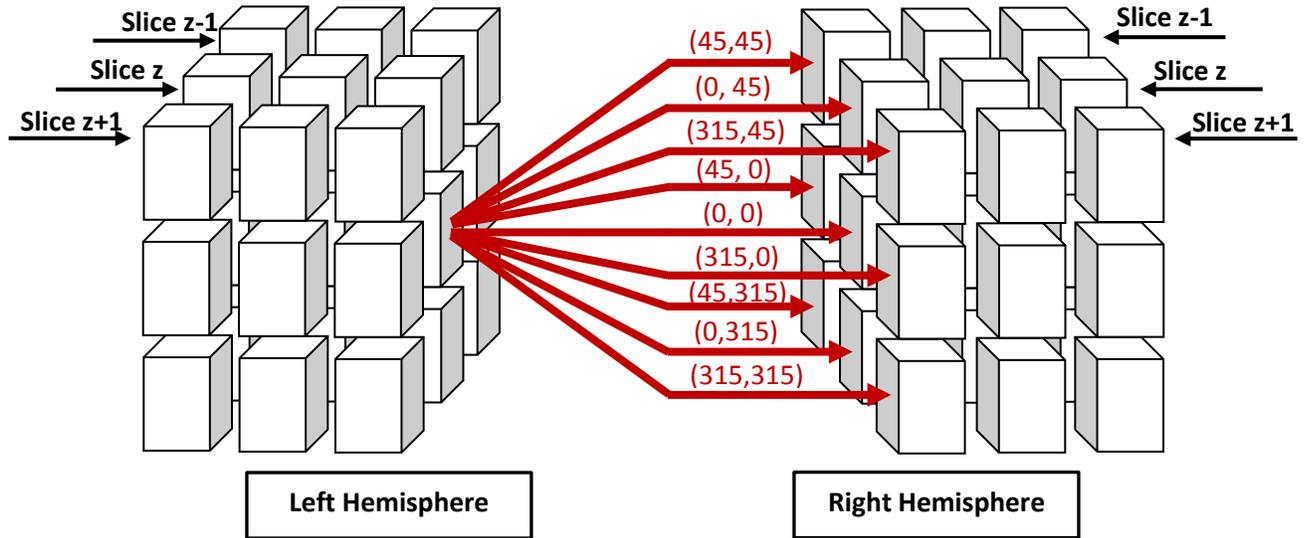


Figure 6.2: How reference pixel relates with opposite nine pixels.

Thereafter, each co-occurrence matrix is normalized by the total number of its elements to calculate the co-occurrence relative frequency between the grey levels of joint pixels in the brain hemispheres. The nine co-occurrence matrices are defined using Eq. 6.1.

$$P(i,j)_{(\theta_1,\theta_2)} = \frac{1}{256^2} \sum_{x=1}^{512} \sum_{y=1}^{256} \sum_{z=1}^Z \begin{cases} 1, & \text{if } L(x,y,z) = i \text{ and } R(x + \Delta x, y + \Delta y, z + \Delta z) = j \\ 0, & \text{Otherwise} \end{cases} \quad 6.1$$

where L and R denote the left and right volumetric hemispheres respectively, both of them have size of $(512 \times 256 \times z)$ pixels. z represents the number of MRI slices of scan. P is the resultant co-occurrence matrix. Δx , Δy and Δz are changed upon the directions of measured matrix.

For slice z

If $\theta_1=0$ and $\theta_2=0$ then $\Delta x=0$ and $\Delta y=0$ and $\Delta z=0$,

If $\theta_1=0$ and $\theta_2=45$ then $\Delta x=-1$ and $\Delta y=0$ and $\Delta z=0$,

If $\theta_1=0$ and $\theta_2=315$ then $\Delta x=1$ and $\Delta y=0$ and $\Delta z=0$.

For slice $z+1$

If $\theta_1=45$ and $\theta_2=0$ then $\Delta x=0$ and $\Delta y=0$ and $\Delta z=1$,

If $\theta_1=45$ and $\theta_2=45$ then $\Delta x=-1$ and $\Delta y=0$ and $\Delta z=1$,

If $\theta_1=45$ and $\theta_2=315$ then $\Delta x=1$ and $\Delta y=0$ and $\Delta z=1$.

For slice $z-1$

If $\theta_1=315$ and $\theta_2=45$ then $\Delta x=-1$ and $\Delta y=0$ and $\Delta z=-1$,

If $\theta_1=315$ and $\theta_2=315$ then $\Delta x=1$ and $\Delta y=0$ and $\Delta z=-1$,

If $\theta_1=315$ and $\theta_2=0$ then $\Delta x=0$ and $\Delta y=0$ and $\Delta z=-1$.

The resultant co-occurrence matrices are approximately symmetric around the forward diagonal for patients who have healthy brain, and asymmetrical for pathological patients.

Figure 6.3 shows two examples of abnormal and normal MRI brain scans and corresponding co-occurrence matrix at $\theta_1=0$ and $\theta_2=0$. On the left, the MRI brain scans of two patients who have pathological and normal brain scans respectively. On the right, the corresponding co-occurrence matrices of those patients. It can be clearly seen that the corresponded co-occurrence matrix of pathological brain is wider and asymmetry around the forward diagonal. Whereas it is slightly narrower and symmetry around the forward diagonal of normal patient. The computing time for implementing 3DMGLCM for the patient who has 10 MRI slices is about 920 sec. by using an HP Workstation Z820 with Xeon E5-3.8GHz (Quad-Core), and 16GB of RAM (Random Access Memory). That means, each slice is required $920/10=92$ sec. to determine its co-occurrence matrix.

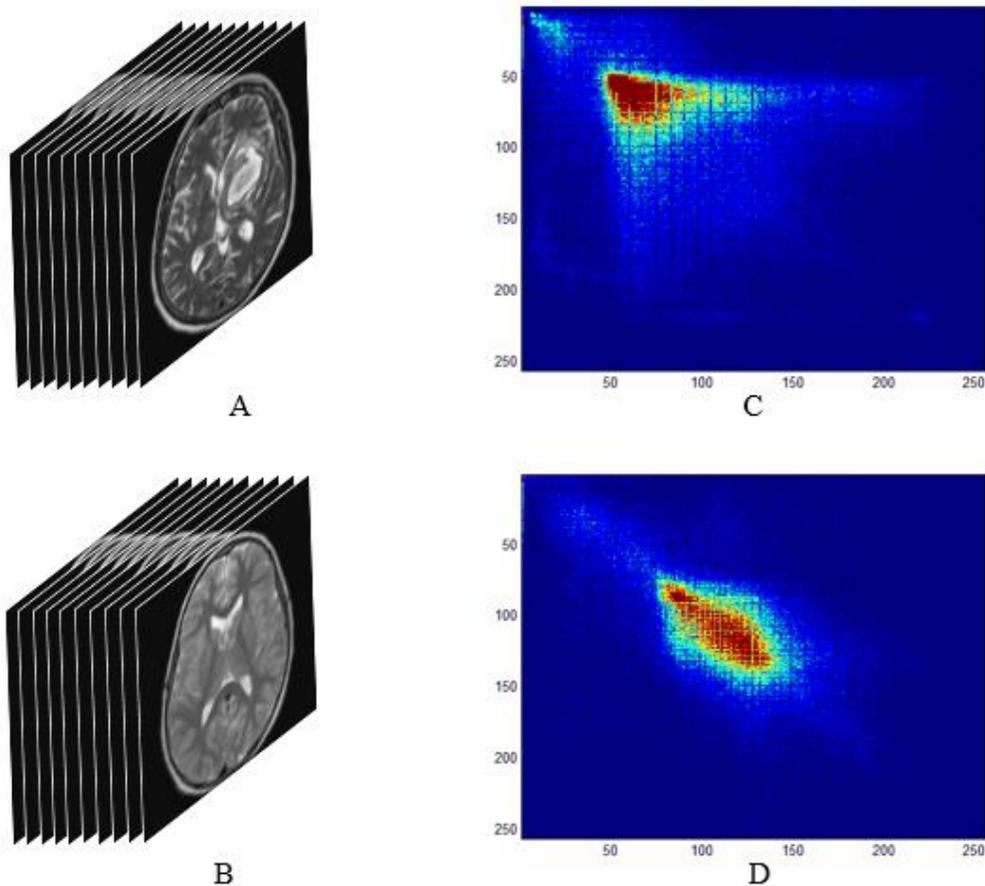


Figure 6.3: A) MRI abnormal brain scan, B) MRI normal brain scan, C) 3DMGLCM of abnormal brain scan and D) 3DMGLCM of normal brain scan.

Finally, 21 co-occurrence statistics that were explained and described in details in section 5.3, are extracted from each co-occurrence matrix. These features are used to measure statistically the degree of symmetry between the two volumetric hemispheres of the brain.

6.3 Feature Aggregation

Due to the proposed 3DMGLCM method determining nine co-occurrence matrices for all MRI slices for each patient, there are 21 descriptors that are determined for each co-occurrence matrix. This mean, there are 189 descriptors for each MRI brain scan. Additionally, the cross-correlation descriptor that is determined for volumetric MRI brain scan. Consequently, there are 190 descriptors for each MRI brain scan.

6.4 Feature Selection

ANOVA method as utilized in the previous chapter in section 5.5, to measure the significance and relevance of features, is used to measure the significance of descriptors which are extracted by 3DMGLCM. As shown in Table 6.1, the differences between the descriptors of normal and abnormal MRI brain scan groups of co-occurrence matrix at $\theta_1=0$ and $\theta_2=0$. It can be clearly seen that some descriptors are highly significant because they have significant P-value, but the degree of significance depends on the F-statistic values. Therefore, different F-statistic threshold values were used to eliminate the descriptors that have small significance as well as minimize the redundancy between features. For instance, once the threshold value is set to 0, the most significant descriptors will be; *contrast*, *correlation*, *dissimilarity*, *energy*, *homogeneity*, *maximum probability*, *difference entropy*, *information measure of correlation I*, *information measure of correlation II*, *inverse difference normalised*, *inverse difference moment normalised* and *cross correlation*.

6.5 Feature Normalization

As explained in section 5.6, all extracted features are normalized by using the min-max normalization approach due to the significant variation between the extracted features that makes some descriptors with larger values influenced more than other descriptors with small values on the behaviour of the classifier (Larose, 2005; Hasan and Meziane, 2016).

6.6 Feature Classification

The same classifiers that were used previously in chapter 5 to classify the extracted features by MGLCM method from MRI brain scans, are used to classify the extracted features by 3DMGLCM. These classifiers are LDA (Zacharaki et al., 2009; Takayanagi et al., 2011), SVM (Nagarajan et al., 2013; Kharrat et al., 2010; Bauer et al., 2011; Hackmack et al., 2012; Kalbkhani et al., 2013) and MLP (Pantelis, 2010; Saritha et al., 2013; Antkowiak, 2006). The training samples are randomly selected and the robustness of the classifiers is validated using the 10-folded cross-validation method. Figure 6.4 shows the flowchart and steps of training and testing of classification techniques. As mentioned in chapter four, the collected dataset includes 165 patients and it was clinically classified into normal and abnormal by the clinicians.

Table 6.1: Comparison of MRI brain scans feature (mean \pm standard deviation (SD)) between normal and abnormal patients.

Features	Abnormal MRI scans	Normal MRI scans	F-statistic	P-value
Auto correlation ($\times 10^3$)	6.3 \pm 1	6.63 \pm 1.2	0.006	0.939
Contrast ($\times 10^3$)	1.96 \pm 0.54	1.243 \pm 0.44	31.578	<0.001
Correlation ($\div 10$)	7.17 \pm 0.72	7.99 \pm 0.73	235.158	<0.001
Cluster Prominence ($\times 10^8$)	3.76 \pm 1.34	3.33 \pm 0.8	1.254	0.266
Cluster Shade ($\times 10^5$)	6.77 \pm 3.29	4.64 \pm 2.9	2.090	0.152
Dissimilarity ($\times 10$)	2.5 \pm 0.43	1.92 \pm 0.34	36.270	<0.001
Energy ($\div 100$)	9 \pm 1.8	8.8 \pm 1.9	328.162	<0.001
Entropy	7.4 \pm 0.29	7.35 \pm 0.3	11.069	0.001
Homogeneity ($\div 10$)	3.37 \pm 0.315	3.44 \pm 0.311	757.204	<0.001
Max. Probability ($\div 10$)	2.99 \pm 0.318	2.95 \pm 0.325	689.810	<0.001
Sum of Square Variance ($\times 10^3$)	7.12 \pm 1.2	7.27 \pm 1.2	0.424	0.517
Sum Average ($\times 10^2$)	1.24 \pm 0.1	1.27 \pm 0.12	0.250	0.619
Sum Variance ($\times 10^4$)	2.6 \pm 3.99	2.66 \pm 4.6	0.360	0.55
Sum Entropy	4.63 \pm 0.15	4.61 \pm 0.158	0.108	0.744
Difference Entropy	3.74 \pm 0.17	3.55 \pm 0.17	21.444	<0.001
Information Measure of Correlation I ($\div 10$)	-1.8 \pm 0.23	-2.06 \pm 0.26	561.403	<0.001
Information Measure of Correlation II ($\div 10$)	8.78 \pm 0.225	9.02 \pm 0.2	378.739	<0.001
Inverse Difference Normalized ($\div 10$)	9.22 \pm 0.11	9.38 \pm 0.092	485.585	<0.001
Inverse Difference Moment Normalized ($\div 10$)	9.75 \pm 0.064	9.83 \pm 0.052	232.709	<0.001
Weighted Mean	-5.96 \pm 28.7	0.57 \pm 6.36	0.726	0.397
Weighted Distance	3.2 \pm 3.02	0.8 \pm 0.58	12.733	0.001
Cross Correlation ($\div 10$)	7.17 \pm 0.72	7.99 \pm 0.37	235.158	<0.001

The average performance of classifying the MRI brain scans into normal and abnormal scans by LDA and MLP are 73.6% and 87.68% respectively. While the highest achieved accuracy with best performance was achieved by SVM with 88.2%. The performance of 3DMGLCM method was compared with the achieved results by MGLCM, GLCM and Gabor wavelet methods that were demonstrated in chapter 5. It is noted that there is a superiority in classification accuracies of MGLCM than others in LDA and MLP. While, 3DMGLCM outweighed MGLCM and GLCM when using SVM classifier as shown in Fig. 6.5.

After implementing ANOVA method for relevance analysis, different F-statistic threshold values were tested at each run and the results of the classification are shown in Table 6.2.

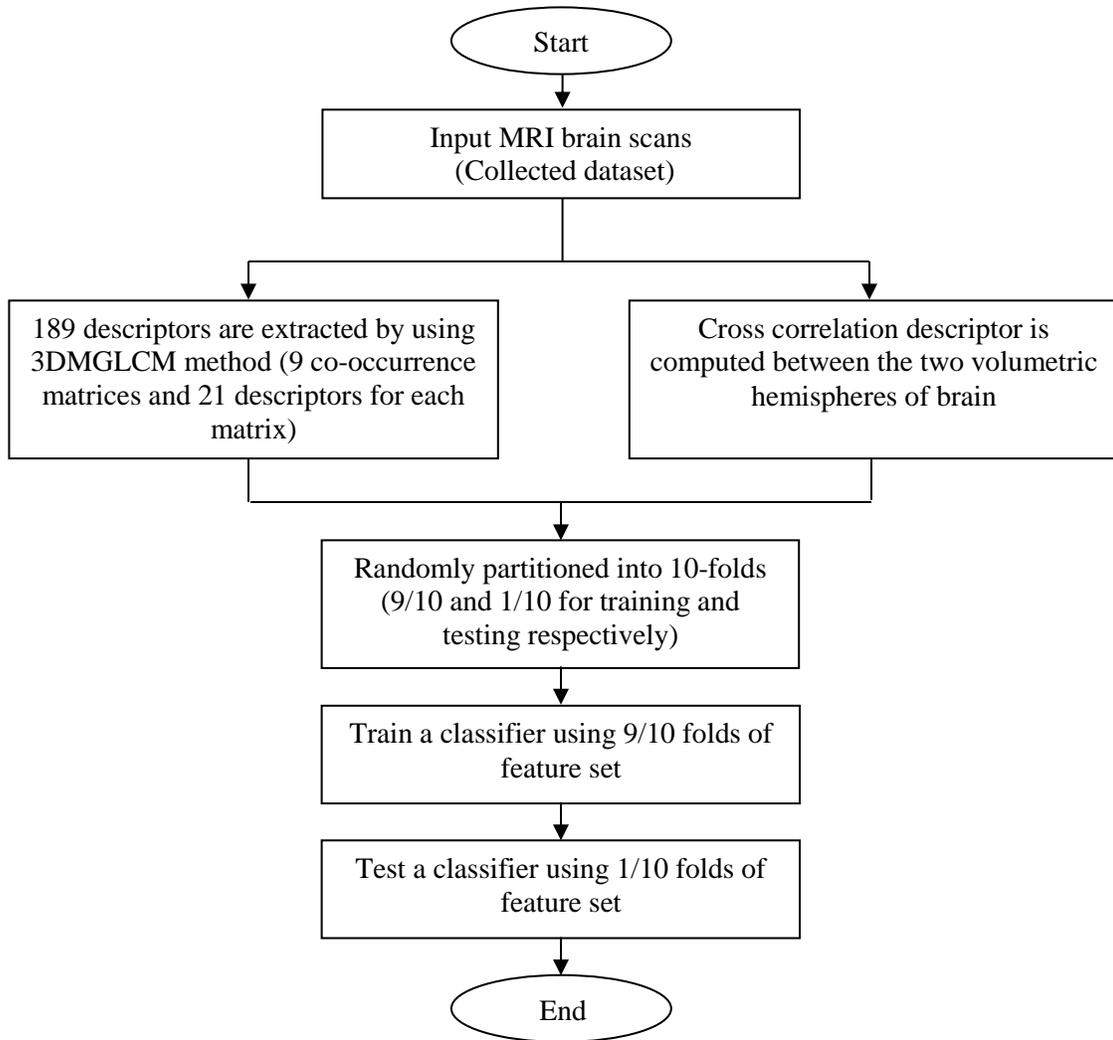


Figure 6.4: Flowchart and steps of training and testing of classification techniques.

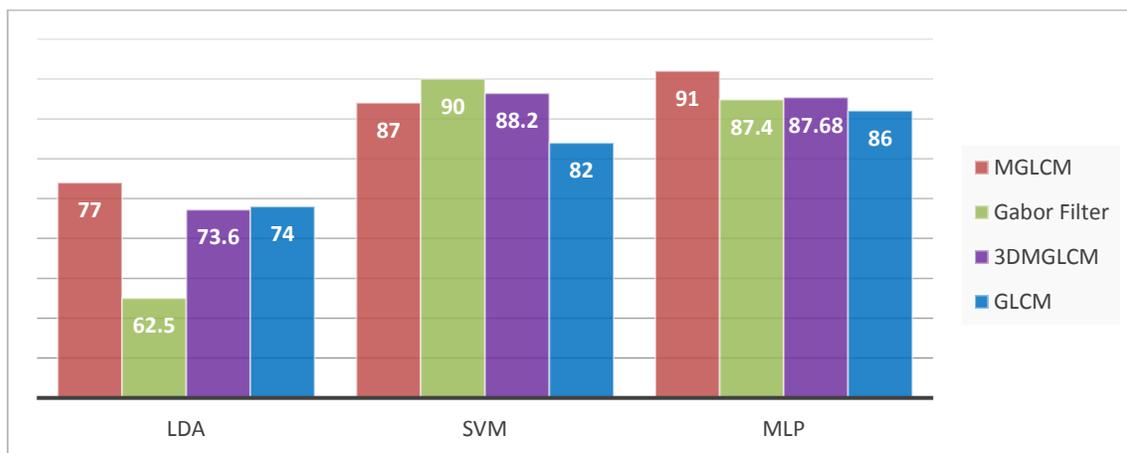


Figure 6.5: A comparison of classification accuracies for 3DMGLCM, MGLCM, Gabor wavelet and GLCM.

Table 6.2: The result of using four different F-statistic threshold values on the classification accuracy.

F-statistic threshold value	No. of selected descriptors	LDA Accuracy	SVM Accuracy	MLP	
				Neurons in hidden layer	Accuracy
0	100	76.8%	88.2%	100	93.3%
30	91	77.3%	88.2%	50	91.8%
200	73	61.2%	89.3%	60	87.2%

It is noted that the classification accuracies of the three classifiers have improved and the best performance of MLP was 93.3% at 0 threshold value, compared to SVM was 89.3% at 200 threshold value and 77.3% was achieved by LDA at 30 threshold value as shown in Fig. 6.6.

The transfer function that was used in the MLP network was the tangent function. To update the weights and bias value, the scaled conjugate gradient method (*trainscg*) in MATLAB R2013a Image Processing Toolkit (Matlab, 2013), was used to be faster than the default function (*trainlm*) for larger datasets. In order to decide the number of neurons in the hidden layer, 10 runs with different number of neurons in the hidden layer are implemented and the RMSEs of runs are shown in Table 6.3.

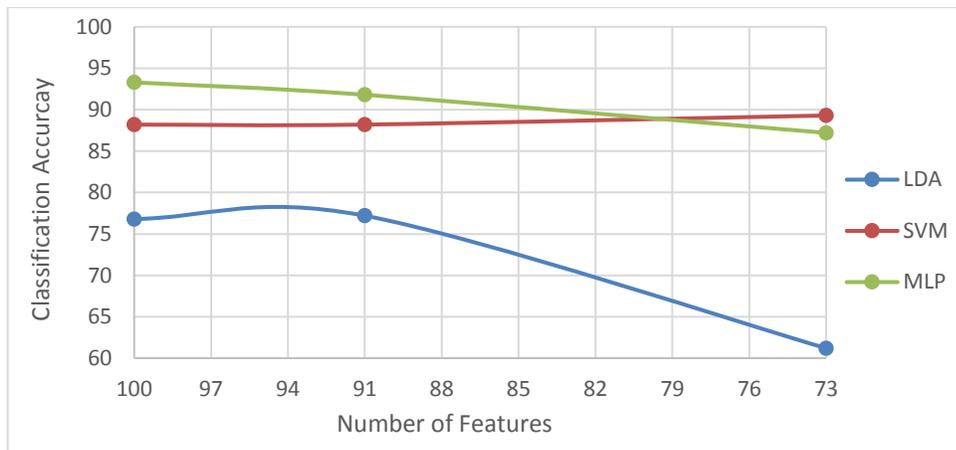


Figure 6.6: The optimal number of features corresponding with classification accuracy.

Table 6.3: The results of 10 runs for 100 input descriptors with different number of neurons in hidden layer.

Neurons No.	RMSE of Runs										Average
	1	2	3	4	5	6	7	8	9	10	
50	0.463	0.422	0.422	0.463	0.378	0.5	0.5	0.415	0.422	0.5	0.448
70	0.463	0.5	0.463	0.327	0.463	0.415	0.463	0.422	0.463	0.378	0.435
80	0.5	0.267	0.267	0.422	0	0.378	0.267	0.327	0.378	0.267	0.307
100	0.422	0.185	0.267	0.327	0.185	0.262	0.189	0.189	0.267	0.189	0.248
110	0.5	0.327	0.267	0.327	0	0.327	0.372	0.378	0.378	9.267	0.31

Figure 6.7 shows the average values of RMSE for runs and it can be clearly seen that the minimum value of RMSE occurred at 100 neurons in hidden layer with average value of 0.248.

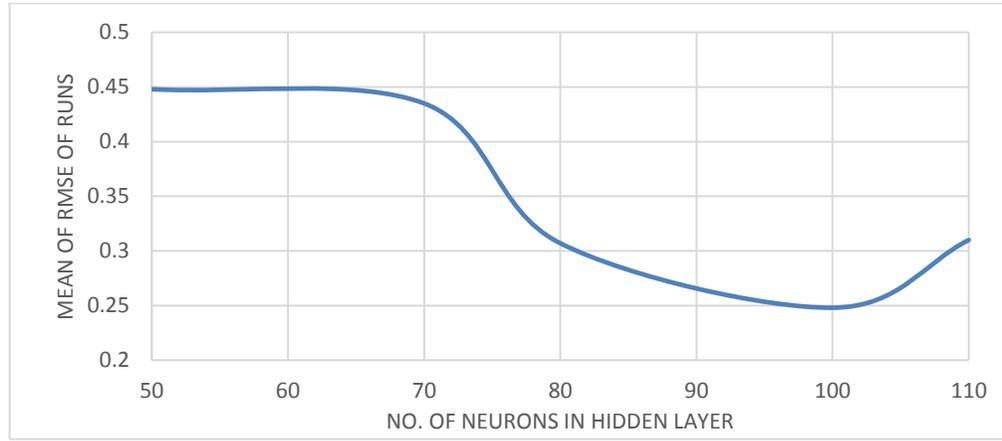
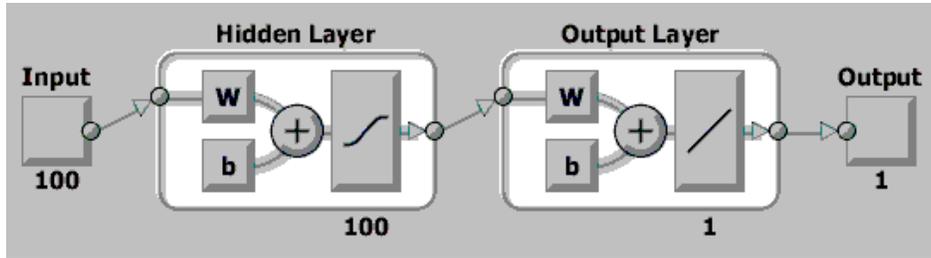


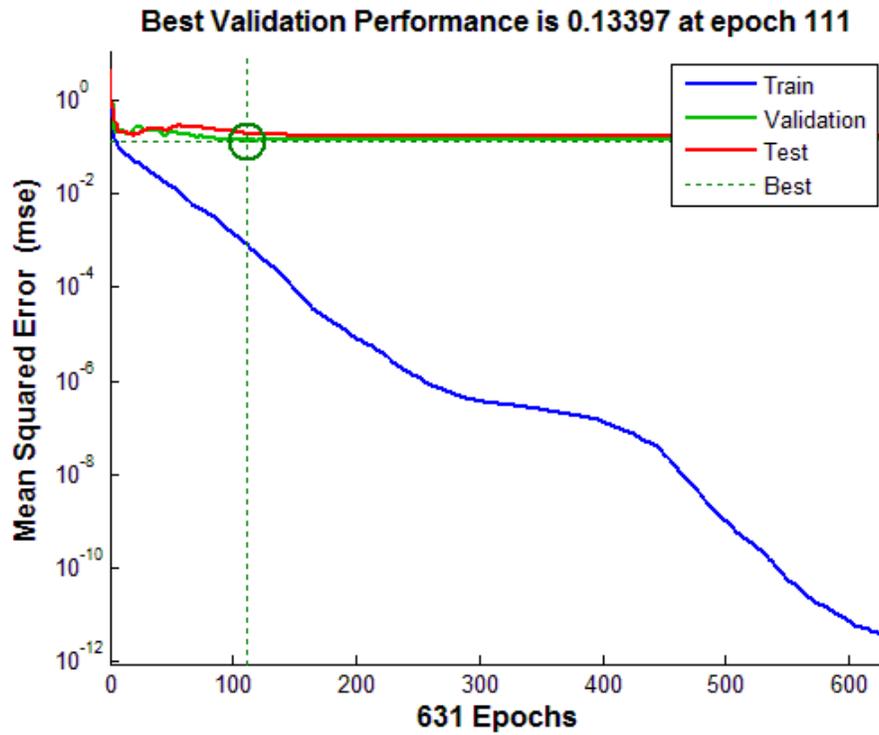
Figure 6.7: Number of neurons vs. RMSEs of runs.

The training of the MLP network with 100 input layer neurons, 100 hidden layer neurons and single output layer is shown in Fig. 6.8. The average performance of the MLP network was $93.3 \pm 0.15\%$. The sensitivity and specificity rates were $94.1 \pm 0.3\%$ and $92.8 \pm 0.3\%$ respectively.

Consequently, the number of descriptors in the feature vector was reduced from 190 to 100. Where, 90 descriptors were discarded and considered as irrelevant or redundant features. Eleven relevant and significant features for each angle of the 3DMGLCM were chosen by the ANOVA method namely: *contrast*, *correlation*, *dissimilarity*, *energy*, *homogeneity*, *maximum probability*, *difference entropy*, *information measure of correlation I*, *information measure of correlation II*, *inverse difference normalized (IDN)*, *inverse difference moment normalized (IDMN)* and *cross correlation*.



A



B

Figure 6.8: A) MLP network structure and B) The performance of MLP network.

Figure 6.9 shows how these selected descriptors are significantly different in means and standard errors (SE) between these two groups.

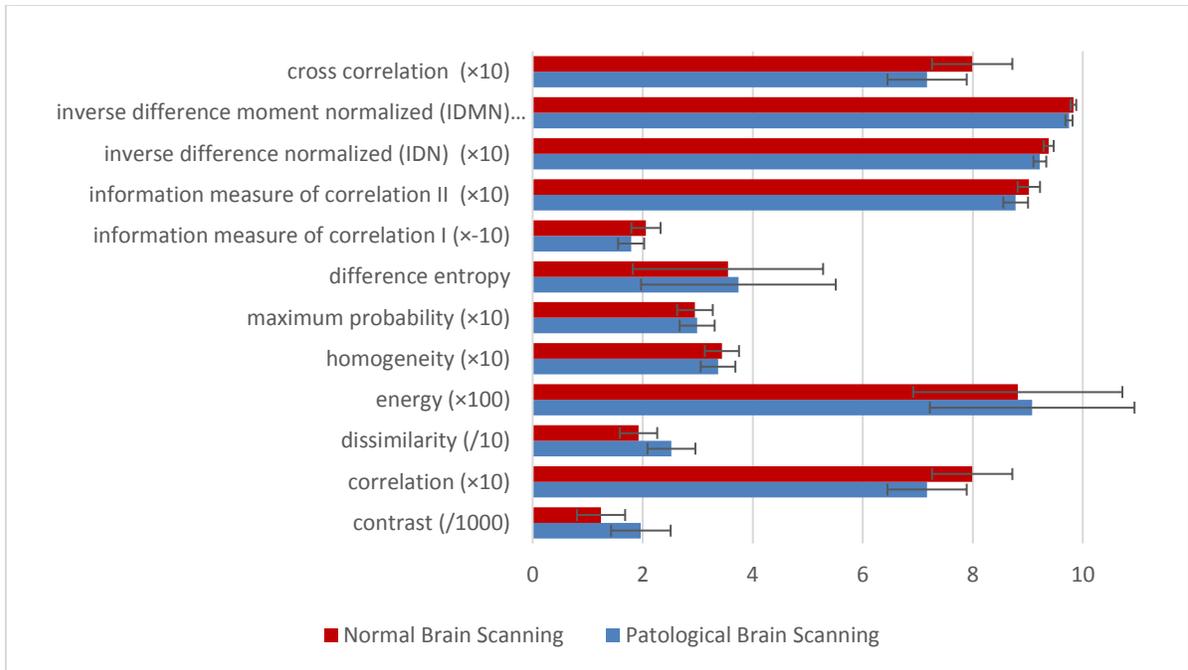


Figure 6.9: Textural features (mean ± standard error (SE)) of the normal and pathological MRI brain scans.

6.7 Conclusion

In this chapter, a novel 3DMGLCM method was implemented on volumetric data of MRI scans instead of extracting texture features from each MRI slice separately with low computational complexity. Over the entire collected dataset and standard dataset (BRATS 2013), the average achieved accuracies by 3DMGLCM were 93.30% and 95.30% respectively by using MLP classifier. Consequently, the computation time for extracting texture features from MRI brain scans is reduced by 38% based on 3DMGLCM method using a single MRI scan modality (T2-weighted image). The 3DMGLCM gives high performance and accuracy in discriminating the normality and abnormality of the brain scans. However, this method has one disadvantage here regarding the high memory requirements.

The identification of the location of the abnormality in MRI brain scan will be investigated in the next chapter.

CHAPTER SEVEN

Detection of the Tumour Location and Slices

Overview

This chapter includes an explanation in details of the proposed method to identify brain tumours location and tumour slices detection automatically based on GA. The proposed method works in a three-dimensional space and does not need for skull elimination.

7.1 Introduction

Automated detection of abnormalities in MRI brain scans is important and necessary in medical diagnostics, planning and treatment. It is more complex than other image objects recognition due to the brain tumours not having regular shapes and standard properties. The general property of a healthy brain is the approximate left-right symmetry (Dvořák et al., 2013). This merit is investigated in this study to detect the location of abnormalities in the brain in order to be able to initiate the segmentation algorithm automatically. Many tumour segmentation methods are not fully automated. These approaches require user involvement in selecting a seed point. Since the main factor in detecting tumours from healthy tissues is the difference in intensity level, such that the tumour appears brighter than the surrounding brain tissue. Tumours are more condensed than the surrounding material and present as brighter pixels than the surrounding brain tissue. Therefore, the basic concept of brain tumour detection algorithms is finding pixel clusters with a different or a higher intensity than that of their surroundings (Khandani et al., 2009). In this study, a bounding 3D-boxes based genetic algorithm (BBBGA) (Hasan et al., 2016a; Hasan et al., 2016b) was introduced for locating the clusters of brain tumours automatically without user interaction. Then the centre of this region can be used as a seed point for initializing the segmentation algorithm.

7.2 Overview of Genetic Algorithms

In the early 1970s, John Holland one of the founders of evolutionary computation, introduced the concept of GA (Haupt and Haupt, 2004). His aim was to make computers emulates what

nature does (Negnevitsky, 2005). The GA is an efficient, adaptive and robust optimization and stochastic search technique based on biological evolution. It is particularly suited for applications involving search and optimization where the space is huge, complex and multimodal (Bandyopadhyay and Pal, 2007). It is similar to other optimization techniques by defining optimization variables, cost function and termination criterion. GA emulates the principle of biological genetics and the principle of the fittest to guide the search and to solve complex optimization problems. It includes a sequence of procedural steps to move from one population which includes a set of individuals to a new one using natural selection and genetic-inspired techniques known as crossover and mutation. These individuals represent the variable domain of the given problem. Some of these individuals that have a higher probability are more likely to be selected to generate better and better populations from old one. Only those individuals in a population who are better suited to solve complex optimization problems are likely to survive and selected according to their fitness in the problem domain. The fitness function is used to measure the performance of each chromosome in the problem domain, and represents the basis for selecting individuals that will be mate during reproduction (Negnevitsky, 2005). Then they are breed together using GA operators to generate new generation (Chipperfield et al., 1994; Grefenstette, 1986; Negnevitsky, 2005). Figure 7.1 shows the flowchart of a GA (Haupt and Haupt, 2004).

GA is different from other optimization and search techniques in (Grefenstette, 1986):

- 1- GA does not work immediately with the parameters of the problem, but it works with the encoding of these parameters.
- 2- GA works simultaneously in parallel with multiple points, not with a single point therefore it has very little chance to fall at a local minimum when using optimization technique.
- 3- GA is considered a blind search technique because it uses only the payoff information.
- 4- GA searches using stochastic operators instead of deterministic rules.
- 5- GA does not need essentially that the search space is continuous and no auxiliary information is required such as the derivative of the optimizing function.

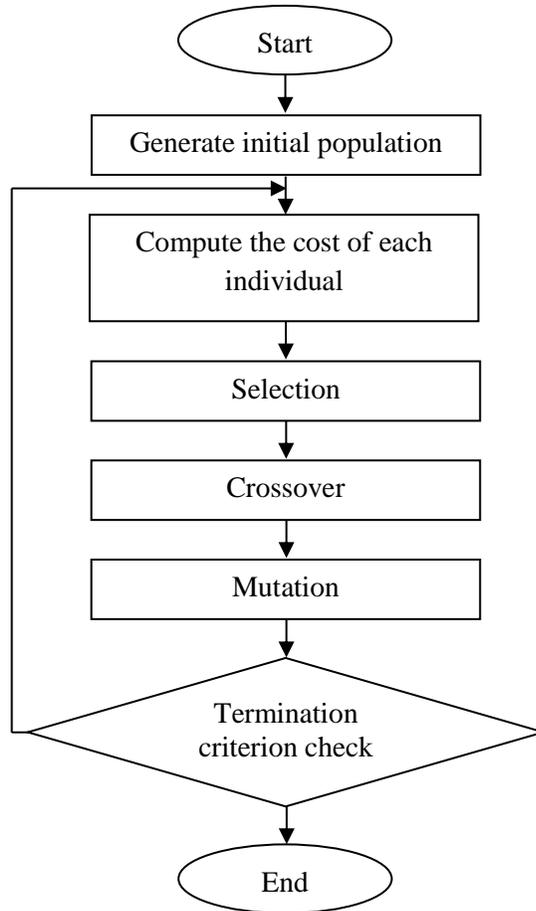


Figure 7.1: Flowchart of GA.

To solve an optimization problem, the GA begins by encoding the individuals as a string, these individuals represent solutions of the problem in the search space. Each of these individuals is comprised of a set of variables or genes to be optimized in a binary form (Chipperfield et al., 1994; Ross, 2009). For instance, if the problem to be solved has two variables x_1 and x_2 , these variables are mapped into the individual structure as shown in Fig 7.2, where x_1 is encoded with 9 bits and x_2 with 12 bits. Now, it is possible to assess and evaluate the performance of each individual in the population by using the fitness function, or it is also known as the objective function.

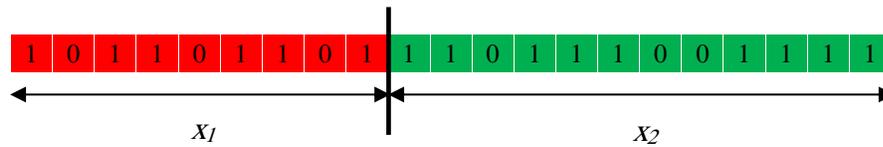


Figure 7.2: How the variables are organized in the chromosome.

After the fitness function is computed for each individual in the population, the survival of the fittest and the death of the poor individual is achieved by applying the three basic operators of GA: selection, crossover and mutation (Talebi et al., 2010; Sonka et al., 2014):

- **Selection operator** is a process of choosing the individuals as parents with the highest fitness from the current population into the new population (Sonka et al., 2014). Then, only the best individuals are selected to continue to the next generation, while the rest is discarded (Haupt and Haupt, 2004). Different selection methods exist for choosing the most fit individuals; roulette wheel selection method, stochastic universal selection method and binary tournament selection method.

Roulette wheel selection method is widely used in many applications of GA. It depends on a probability distribution to choose the fittest individuals. Such that the selection probability of the given individual is directly proportional to its fitness (Colin and Jonathan, 2002). Figure 7.3 shows the roulette wheel of six individual's fitness probability values, individual 2 is the most popular fit and invade the large interval, while individuals 1 and 3 are the least fit and occupy smaller intervals in the roulette wheel. The individual is selected by multiplying the total sum of individual's probability values by a random number that is generated within the interval 0 and 1 (Chipperfield et al., 1994).

Stochastic universal selection method depends on placing P equidistant markers on the wheel as shown in Fig. 7.4, such that all the P individuals are chosen by spinning the wheel and the number of selected individuals are equal to the number of markers that points within the corresponding slot (Bandyopadhyay and Pal, 2007). Practically, instead of a single choice at roulette wheel selection method, this method may be used to choose P individuals at each time (Colin and Jonathan, 2002; Negnevitsky, 2005).

Binary tournament selection method depends on choosing two individuals randomly, and the better of them is chosen as a parent for next generation. This process is repeated till all population is filled (Bandyopadhyay and Pal, 2007). This method is more suitable for larger population sizes because it does not need for sorting as sorting becomes time-consuming for large populations (Haupt and Haupt, 2004). In this study, the roulette wheel selection method will be used in the implementation of GA because it is more popular and efficient in different application (Talebi et al., 2010).

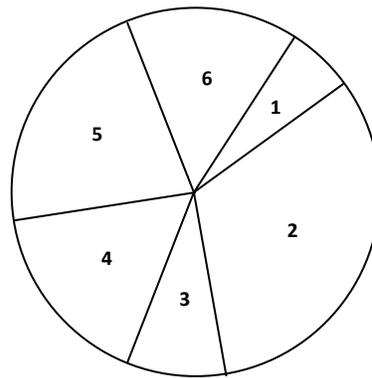


Figure 7.3: Roulette wheel selection method.

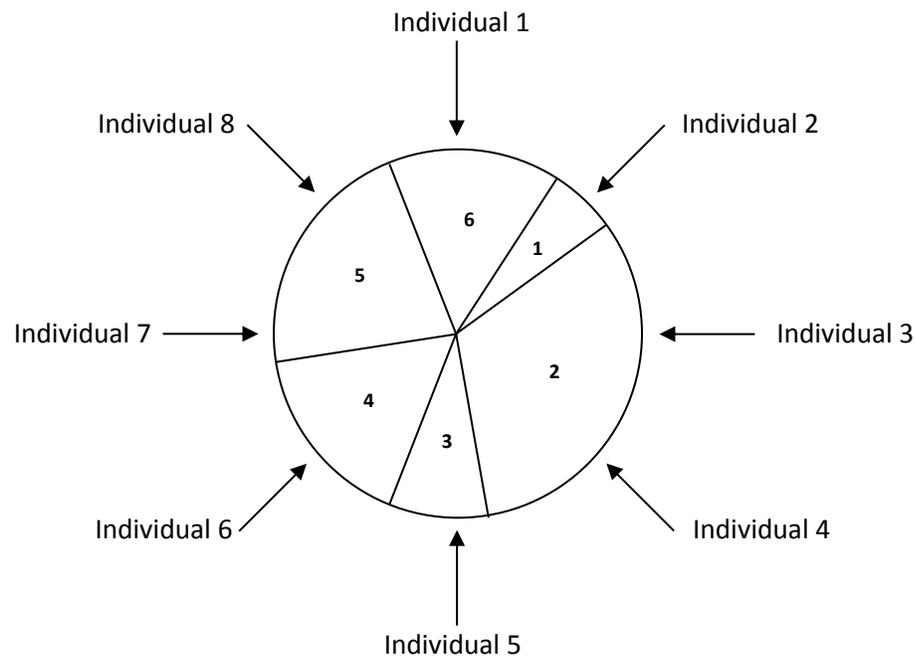


Figure 7.4: Stochastic universal selection method.

• **Crossover operator** is a process of mating the selected parents to exchange information by recombining parts of their genetic materials. The process is performed probabilistically by choosing randomly crossover point and crossing the individuals over to produce new off springs for next generation (Bandyopadhyay and Pal, 2007). The most common form of crossover involves two parents to produce two off springs. The selected pair of individuals undergo crossover with probability P_c . A random number R_c is generated in the range 0 to 1, and compared with P_c . if $R_c \leq P_c$ then the two individuals are undergoing crossover, otherwise they are processed without crossover. Typically the value of P_c is in range 0.4 to 0.9, and if $P_c=0.5$, then half of the new population is formed by selection and crossover, and the other half by selection only (Coley, 1999). There are two types of crossover operators; single point crossover and multi-points crossover. Single point crossover is one of the most commonly used method, it involves generating randomly single crossover point between the first and last bits of the parent's individuals. Such that, the partitions that are located to the right of crossover point for both parents are swapped to produce the first and second offspring (Haupt and Haupt, 2004; Bandyopadhyay and Pal, 2007; Negnevitsky, 2005) as shown in Fig. 7.5.

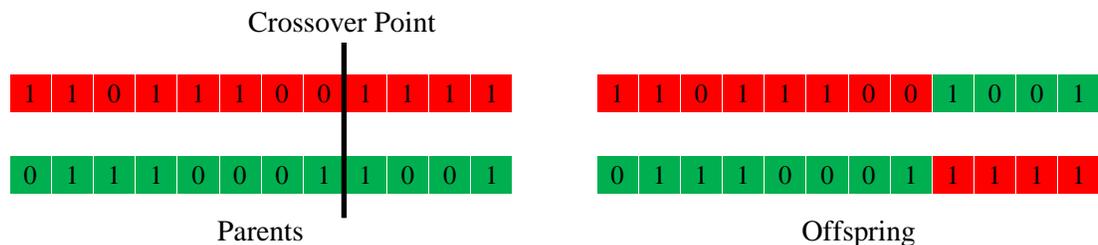


Figure 7.5: Single point crossover operation.

Multi-points crossover involves generating randomly multi-points crossover positions within the length of individual. These multi-points should be without duplication and in an ascending order. Then the parts of individuals between the consecutive crossover points are swapped between the two parents to reproduce two new offspring (Chipperfield et al., 1994) as shown in Fig. 7.6. It will be used in this study within the implementation of GA.

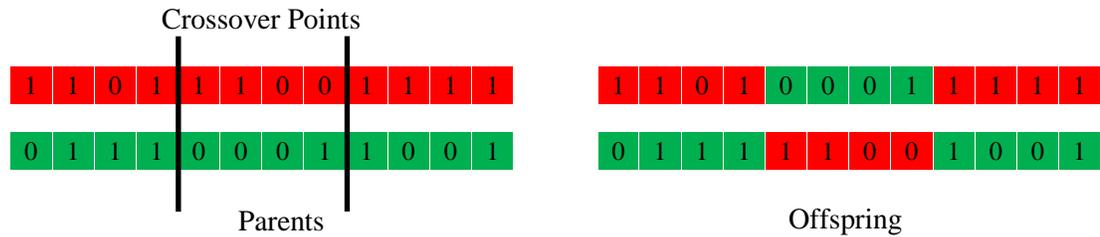


Figure 7.6: Multi-points crossover operation.

- Mutation operator is a process of a random alternation in the genetic structure of an individual to make genetic diversity into the population. It helps to converge to the optimal solution in the search space, if it is not represented in the population (Bandyopadhyay and Pal, 2007). It is implemented by generating a random number within the length of individual and complement the chosen bit according to the random number (Colin and Jonathan, 2002). The probability of mutation P_m is typically of the order 0.001, that means one bit in every thousand will be mutated (Coley, 1999). Similarly, to other genetic operators, there are a single point mutation and multi-points mutation. The single point deals with a single bit by changing a 1 to a 0 and vice versa for each individual as shown in Fig. 7.7. Whereas the multi-points mutation deals with multi-bits for each individual and it helps to increase the algorithm's freedom to search outside the current region of the variable space (Haupt and Haupt, 2004). After implementing all GA operators on the selected parents, the two new offspring are generated and this process is repeated until a new set of N individuals are generated, where N represents the population size. In order to ensure the survival of the best individuals, the bottom half of current population is discarded and replaced by the top half after resorting all the individual descending according to the fitness value (Haupt and Haupt, 2004; Negnevitsky, 2005).

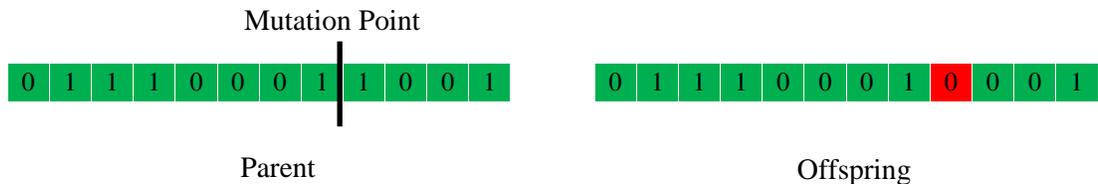


Figure 7.7: Single point mutation operation.

7.3 Termination of GA

Since the GA is considered as a stochastic search method, it is possible to run continuously forever because there are no convergence criteria that can be specified. As well as the fitness value of individuals may remain static for several iterations before the best individual is found. Therefore, there are several criteria that can be used to terminate the GA; specify the maximum number of generations, test the quality of the best individuals of the population against the problem definition and compute the error between current and previous generation. If there is no acceptable solutions may be found, the GA should be restarted (Chipperfield et al., 1994; Colin and Jonathan, 2002; Negnevitsky, 2005).

In this study, two criteria are used to terminate the GA; the maximum number of iterations that is set experimentally at 100 and RMSE between current and previous generations. Such that the RMSE value is compared with threshold value that it was set experimentally to 0.5 which denotes the acceptable approximate solution.

7.4 Bounding 3D-Boxes Based Genetic Algorithm Method

The novel bounding 3D-boxes based genetic algorithm (BBBGA) method is proposed in (Hasan et al., 2016a; Hasan et al., 2016b) to search and identify the location of the most dissimilar regions between the left and right hemispheres of the brain automatically without the need for user interaction. The input is a set of MR slices belonging to the scans of a single patient, and its output is a subset of slices covering and circumscribing the tumour with 3D-box. The BBBGA method exploits the symmetry feature of axial viewing of MRI brain slices to search for the most dissimilar region between the left and right brain hemispheres. This dissimilarity is detected using GA and an objective function based mean intensity computation. The process involves randomly generating hundreds of 3D-boxes with different sizes and locations in the left brain hemisphere. Such boxes are then compared with the corresponding 3D-boxes in the right brain hemisphere through the objective function. These 3D-boxes are moved and updated during the iterations of the GA toward the region that maximized the objective function value. The objective function value is high when the 3D-box stands on the tumour region and low when standing on soft tissue because the tumour is always brighter than the soft surrounding tissue of the brain in T2-w images (Khandani et al., 2009). The BBBGA is applied to a collected

dataset that was already pre-processed, and the output is the MRI slices that contain the tumour and corresponds to the optimized 3D-box that bounded the tumour over the relevant subset of slices. Many studies have investigated this idea of symmetry of human brain for various applications (Saha et al., 2012; Ray et al., 2008b; Dvořák et al., 2013). These studies have exploited the symmetry of the brain for detecting brain tumours and made the detection task robust. An advantage of BBBGA method is its lack of necessity for image registration. The approach is an unsupervised method; hence, the problems on observer variability in supervised techniques are ignored. It can be implemented in real time as well as it can be used as a seed point to initialize the segmentation process.

7.4.1 Exponential Transformation of MRI Brain Slices

Prior to BBBGA, exponential transformation is implemented to compress the low-contrast regions in MRI brain images and expand the high-contrast regions in a nonlinear manner. This action would increase the intensity difference between the brain tumour and the surrounding soft tissue (Khandani et al., 2009) as shown in Fig. 7.8. This will help the GA to converge and move the generated 3D-boxes faster and accurately to the abnormal region of the brain. Figure 7.9 illustrates the pseudo-code for BBBGA

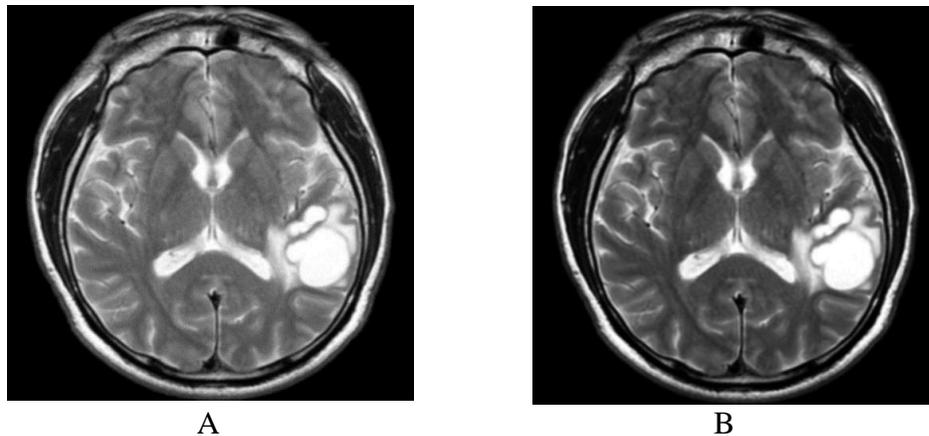


Figure 7.8: Image exponential transformation, A) Original MRI brain slice, B) Transformed MRI brain slice.

```

1. read  $Img \leftarrow$  MRI slices
2.  $Img \leftarrow \exp(Img)$ 
3. // Initialization
4. generate  $N$  feasible individuals  $ind$  randomly in current population
   , which is set experimentally to 100.
5. compute  $mean(Img(ind_i))$  for each  $i \in N$ 
6. // loop until termination condition is achieved
7. for  $i = 1$  to  $N$ 
8. // Selection
9. select the best two individual from current population ( $ind_1, ind_2$ )
10. // Crossover
11.  $newind_1, newind_2 \leftarrow$  with crossover-probability crossover  $ind_1, ind_2$ 
12. // Mutation
13.  $newind_1 \leftarrow$  with mutation-probability mutate  $ind_1$ 
14.  $newind_2 \leftarrow$  with mutation-probability mutate  $ind_2$ 
15. // Evaluation  $newind_1, newind_2$ 
16. compute  $mean(Img(newind_1))$ 
17. compute  $mean(Img(newind_2))$ 
18. new population  $\leftarrow newind_1, newind_2$ 
19. endfor

```

Figure 7.9: Pseudo-code for BBBGA.

7.4.2 Designing of GA

There are several issues involved in designing of GAs such as individual size and population size in addition to choose the most appropriate operations such as selection, crossover and mutation methods.

7.4.2.1 Individual Construction

For additional details on how each individual in the GA population is mapped into binary form, we use the following scenario. Suppose we have a MRI brain scan (dimensions $512 \times 512 \times 32$ pixels) of a pathological patient, each individual in the GA population is denoted by the binary representation of the coordinates of one 3D-box (x_1, x_2, y_1, y_2, z_1 and z_2). In this case, x_1 and x_2 represent the height of the 3D-box and are subject to the constraints $1 \leq x_1 < 512$ and $x_1 < x_2 \leq 512$. Meanwhile, y_1 and y_2 signify the width of the 3D-box and are subject to the constraints $1 \leq y_1 < 256$ and $y_1 < y_2 \leq 256$. Finally, z_1 and z_2 represent the depth of the 3D-box and are subject to the constraints $1 \leq z_1 < 32$ and $z_1 < z_2 \leq 32$. Herein, we assume that the maximum number of

MRI slices is 32. Figure 7.10 shows the original generated 3D-boxes by GA, such that each generated red 3D-box in the left brain hemisphere has a corresponding yellow 3D-box in right brain hemisphere. Figure 7.11 shows an example of how the coordinates of 3D-box $(x_1, x_2, y_1, y_2, z_1, z_2)$ are mapped to the individual of the GA in a binary form.

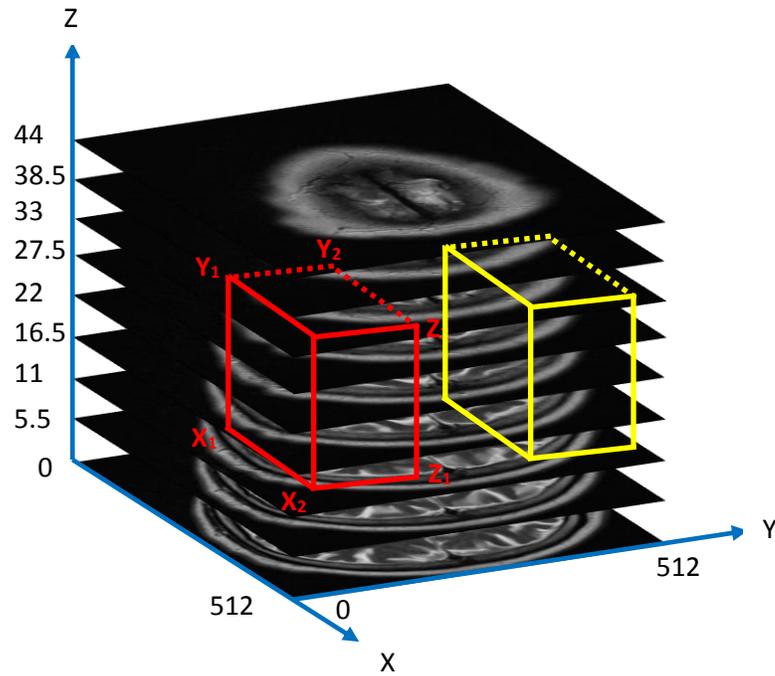


Figure 7.10: Representation of one 3D-box in the left hemisphere of brain using $(x_1, x_2, y_1, y_2, z_1, z_2)$ coordinates and opposite region.

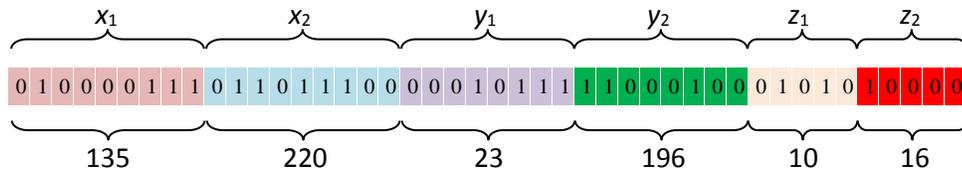


Figure 7.11: Individual structure.

Consequently, the individual size becomes equal to 44 bits. By using the objective function, we can measure the performance of individuals in the problem domain (Chipperfield et al., 1994). In this study, the fittest individuals that have the highest numerical value of the associated objective function are preserved.

The objective function g that is used in this study is based on finding the absolute value of subtracting the means of the intensities inside the generated 3D-box in the left hemisphere from the corresponding 3D-box in the right hemisphere using Eq. 7.1.

$$g = \frac{1}{x, y, z} \left| \sum_{i,j,k}^{x,y,z} I_L(i, j, k) - \sum_{i,j,k}^{x,y,z} I_R(i, j, k) \right| \quad 7.1$$

where x, y and z are the coordinates of the generated 3D-box on the left hemisphere I_L and the corresponding opposite 3D-box in right hemisphere I_R .

7.4.2.2 Population Size

The choice of the population size is considered as an important issue in GA applications because it affects both the ultimate performance and the efficiency of the GA. However, a too small population size would be insufficient for exploring the entire search space effectively and a too large population size would be efficient but it requires a significant amount of computations and more evaluations per generation. This leads to possibly resulting in an unacceptable slow rate of convergence (Colin and Jonathan, 2002). Grefenstette (1986) suggested that the adequate population size is as small as 30 in many cases and the best off-line GA had population size equal to 80 (Haupt and Haupt, 2004). While Haupt (2000) reported that there is a relationship between the population size and mutation rate and the choice of them can impact on the run time and ability of GA to converge to optimal solution.

Table 7.1 illustrates the average number of iterations to converge to the optimal solution of locating the brain tumour in MRI brain scans of 40 patients, by running the BBBGA, 5 times for different population sizes ranging from 20 to 100 in increments of 20, and 5 different mutation rates ranging from 0.05 to 0.25 in increments of 0.05. It is noted that the minimum

average number of iterations was achieved by setting population size to 100 and mutation rate to 0.05 as shown in Fig. 7.12.

Table 7.1: An average number of iterations with different population size and mutation rates.

Mutation rate	Population size				
	20	40	60	80	100
0.05	85	52	44	29	18
0.1	92	72	50	45	36
0.15	99	77	68	47	43
0.2	124	84	76	62	50
0.25	130	104	79	67	74

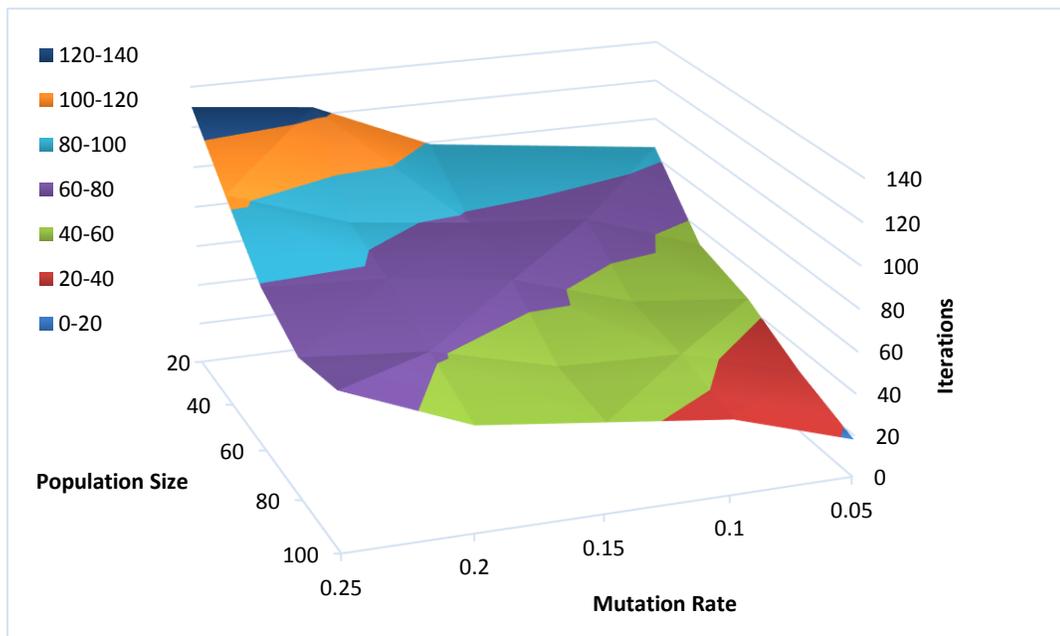


Figure 7.12: Average of GA iterations for different values of population size and mutation rate.

After the MRI brain scans are classified into normal and pathological images, the BBBGA method was applied on those identified as pathological cases as shown in the pathological patient in Fig. 7.13. The red rectangles denoting the optimized 3D-box refer to the pathological area in slices 6–9 where the tumour appears. Figure 7.14 shows how the RMSE decreases significantly over 18 iterations by GA.

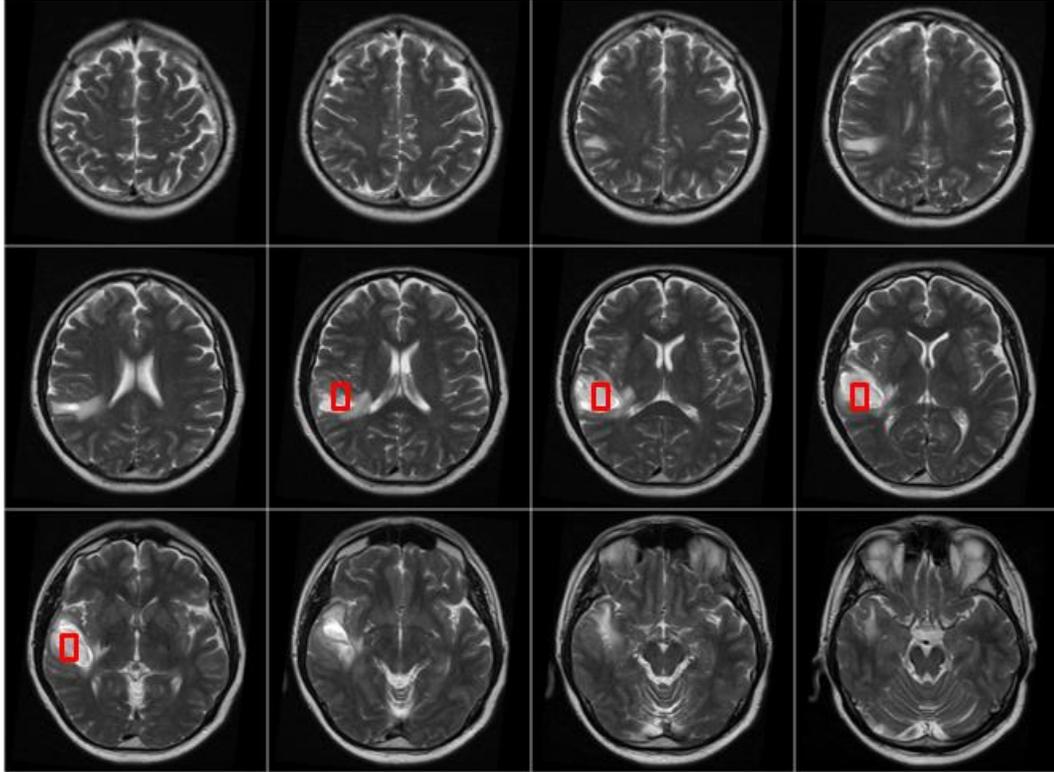


Figure 7.13: MRI brain slices: the output of the BBBGA implementation on pathological MRI brain slices with a population size equal to 100 and a mutation rate equal to 0.05. The red rectangles that denotes the optimized 3D-box stands on pathological area in slices 6 to 9, where the tumour are appeared.

The BBBGA method was implemented on MRI brain scans of pathological patients with population size (N) equal to 100. The individuals were selected using the roulette wheel selection method because this approach is more popular and efficient in different applications (Talebi et al., 2010). The selected individuals were then mated using a multi-point crossover with probability of 0.5 (Chipperfield et al., 1994). Finally, a single-point mutation was implemented with a probability of 0.05.

Evidence extracted from previous studies (Khandani et al., 2009; Saha et al., 2012) indicates the lack of a standard method for evaluating the BBBGA method. Saha et al. (2012) used an example to observe and measure the noise sensitivity of his approach by adding Gaussian noise with different values of $\sigma = 0, 0.1, 0.2, 0.3$ and 0.4 although this addition is not important and irrelevant for the evaluation.

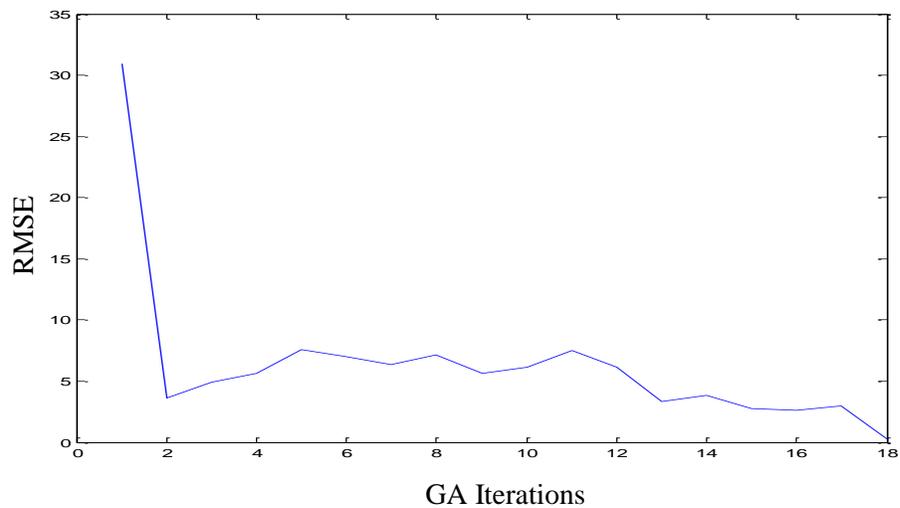


Figure 7.14: RMSE decreases within 18 iterations by GA.

The noise sensitivity of BBBGA is measured after the addition of the Gaussian noise of the same noise amounts and this is shown in Appendix A. Figure 7.15 shows that FP is proportional to the amount of noise in the MRI scan. Hence, our approach was evaluated using the collected dataset that included 88 pathological cases. Among 84 pathological cases, an abnormality was successfully located. Only four cases remained undetected because of the method's inability to detect hardly visible tumours of size less than 1 cm^3 . Moreover, tiny tumours hold a spatial scale relatively similar to normal anatomic variability (Sanjuán et al., 2013).

Table 7.2 shows the comparison of identifying clinically and experimentally the most important slices of 12 pathological patients. It is noted, that the BBBGA method has succeeded to identify number of pathological MRI slices for each patient. The located pathological MRI slices are within the range of important slices that were identified clinically.

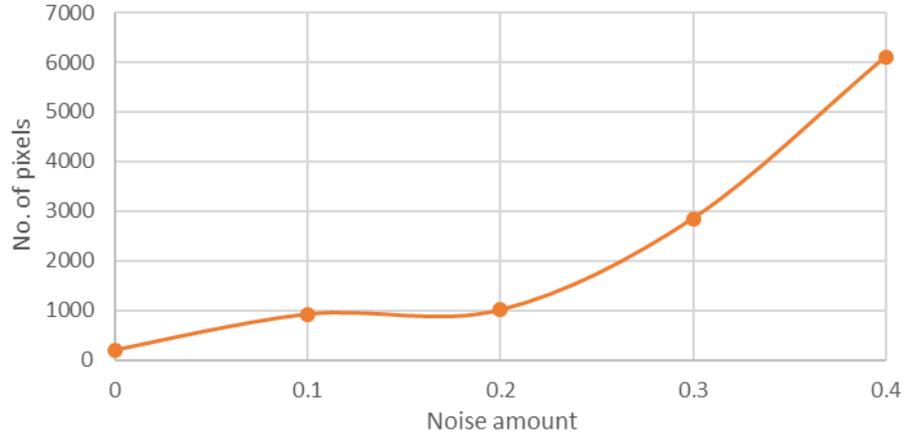


Figure 7.15: FP increases with increasing noise amount in the MRI scan.

Table 7.2: A comparison between clinical and experimental slice identification.

Patient No.	Clinically Identified		Experimentally Identified	
	Tumour Starting Slice	Tumour Ending Slice	Tumour Starting Slice	Tumour Ending Slice
1	5	11	5	8
2	4	11	7	9
3	3	11	9	10
4	3	9	3	7
5	2	12	9	11
6	5	14	5	12
7	8	13	9	13
8	16	22	20	22
9	6	10	8	10
10	4	10	8	9
11	2	11	9	10
12	1	13	6	8

7.5 Conclusion

In this chapter, the novel BBBGA method was presented for localizing brain tumours and recognizing the most relevant pathological slices that will draw the attention of clinicians to diagnose these slices immediately without requiring the inspection of all the patients' slices. It exploits the symmetry feature of axial viewing of MRI brain slices (T2-w images) to search for the most dissymmetric region in the volume of the brain using GA and region score function that uses the mean of intensity. It is a completely unsupervised method that does not require a training phase and does not need for image registration, intensity standardization and skull

removing. The achieved accuracy by the BBBGA to locate the brain tumours was 95%, with only 4 cases out of 88 pathological patients whose tumour's size is less than 1 cm³, have failed to be identified the abnormality because the relative similarity of spatial scale between the tiny brain tumours and normal variability in anatomy (Hasan et al., 2016b). Consequently, the proposed BBBGA has achieved acceptable result in indexing brain tumours that make it eligible to initialize the seed point of segmentation process as it will be demonstrated in next chapter.

CHAPTER EIGHT

Brain Tumour Segmentation

Overview

This chapter demonstrates the results of the brain tumours segmentation that are obtained from implementing 3DACWE in addition to identifying the most important MRI slices that present brain tumour. It also includes a comparison between the achieved results by 3DACWE and 2DACWE, and other proposed techniques for segmenting brain tumours.

8.1 Introduction

In the medical diagnostics field, brain tumours are relatively less common compared to other neoplasms such as lung and breast, but are considered highly important due to their effect on the patient and present a high morbidity (Karkavelas and Tascos, 2011). They require specific studies because of the complicated pathologies which make them difficult to diagnose. Accurate detection and segmentation of brain tumours have a significant influence on clinical diagnosis, predicting prognosis and treatment. In addition, they are beneficial for general modelling of pathological brains and the anatomical construction of the brain (Nabizadeh and Kubat, 2015). In section 2.6.4, the deformable model has been proved to be efficient. It is one of the most recommended method that have been used for different applications of medical images segmentation due to its capability of accommodating the often significant variability of biological structures over time and cross different patients (Rousseau, 2009; Tantisatirapong, 2015).

An evaluation procedure is important to estimate the reliability and quality of segmentation techniques. Image segmentation evaluation can be categorized into subjective and objective evaluations. The subjective evaluation method requires to compare visually the result of the image segmentation with one or more human experts. Since each human expert has his own experience and distinct standards for evaluating and assessing the segmentation results, the evaluation results may significantly differ from one expert to another. Consequently, the

subjective evaluation is considered a very tedious procedure and time consuming (Zhang et al., 2008). While the objective evaluation is divided into supervised and unsupervised techniques. Supervised evaluation methods evaluate segmentation algorithms by comparing segmentation results with manually-segmented reference images which are segmented by experts and reflects the optimum of the resulting segmentation. It is also known as ground truth reference images or gold standard. The degree of similarity between the human and machine segmented images determines the quality of the segmentation (Zhang et al., 2008). While unsupervised evaluation methods do not require to compare with any additional reference images, it just relies on the degree of matching among the characteristics of segmented images as desired by humans. The main advantage of unsupervised evaluation method is that it does not need to compare against a manual segmented reference image. This merit makes it more suitable for real-time application where a large number of images with unknown content and no ground truth need (Tantisatirapong, 2015; Beneš and Zitová, 2015; Menze et al., 2015; Christos, 2005; Zhang et al., 2008).

In this study, supervised evaluation is preferred because of the complexity of the brain tissue and variety of brain tumours as well as its ability to distinguish slight differences between the outputs of different segmentation techniques (Beneš and Zitová, 2015; Zhang et al., 2008). This evaluation method measures the degree of similarity between the segmented brain tumours and those that are segmented manually or with a ground truth dataset. A set of statistical measures have been used to evaluate the segmentation outcomes. They are TP, FP, TN and FN. The TP denotes the number of pixels that are correctly segmented as part of a tumour, FP denotes the number of pixels that are incorrectly segmented as part of a tumour, TN denotes the number of pixels that are correctly segmented as a healthy pixels and FN denotes the number of pixels that are incorrectly segmented as a healthy pixels as shown in Fig. 8.1 (Anbeek et al., 2005; Tantisatirapong, 2015). These measures are used to evaluate the segmentation process; accuracy, sensitivity and specificity. Accuracy is defined as the ratio of numbers of pixels that are correctly segmented to the total number of pixels in MRI slices and given in Eq. 5.41. Sensitivity considers the proportion of the tumour that is correctly segmented and given in Eq. 5.42. Specificity refers to non-tumour portion that is correctly segmented and given in Eq. 5.43. (Nabizadeh, 2015; Udomchaiporn et al., 2013; Menze et al., 2015).

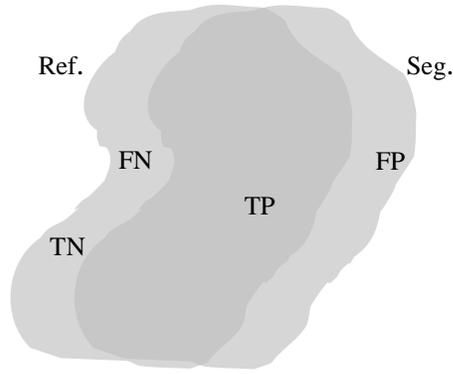


Figure 8.1: Comparison of a binary segmentation image (Seg.) with the reference image (Ref.) and the statistical measures.

In the evaluation of the segmentation, both accuracy and specificity are not highly relevant because these two measures adopt the TN, which depends on the relative size of the MRI slices. Therefore, the following additional metrics are used to evaluate the segmentation results: Dice, Jaccard and matching coefficients. The Dice coefficient is used to measure the percentage of spatial overlapping between two binary images. It is more popular in reporting the performance of segmentation. The mathematical representation of the Dice coefficient is given in Eq. 8.1 (Agrawal and Sharma, 2014).

$$Dice\ coefficient = \frac{2|Seg \cap Ref|}{|Seg| + |Ref|} = \frac{2\ TP}{2\ TP + FP + FN} \quad 8.1$$

Jaccard coefficient is another popular measure that is also used to measure the percentage of overlapping between two binary images. The mathematical representation of the Jaccard coefficient is given in Eq. 8.2 (Agrawal and Sharma, 2014).

$$Jaccard\ coefficient = \frac{|Ref \cap Seg|}{|Ref \cup Seg|} = \frac{TP}{TP + FP + FN} \quad 8.2$$

Matching coefficient is used to measure the percentage of matching between two binary images. It is calculated as the ratio of the TP to the number of ground truth tumour pixels as given in Eq. 8.3 (Fletcher-Heath et al., 2001).

$$Matching\ coefficient = \frac{TP}{TP + FN} \quad 8.3$$

The ACWE segmentation method was combined with BBBGA to implement it automatically to improve the capture range of brain tumour boundary in volumetric MRI scans. This method is applied in two and three-dimensional spaces on four modalities of MRI: T2-w, T1-w, FLAIR and T1c-w images.

8.2 Brain Tumour Segmentation

This section includes implementation of 2DACWE and 3DACWE segmentation methods of brain tumours and their evaluation by comparing them with manual segmentation (clinical segmentation) and standard dataset (BRATS 2013). To fully segment the brain tumours, the segmentation algorithm is applied on MRI slices either on slice by slice using 2DACWE or on volumetric MRI scans using 3DACWE. The implementation of these two methods will be described in this chapter using different examples. Finally, a conclusion and a comparison between these two methods will be performed. We first set the initial condition of the level set function \emptyset using the BBBGA method as illustrated in chapter seven to choose a seed point to start the segmentation process. The outputs of the brain tumour segmentation method were compared with 25 patients in the standard dataset BRATS 2013 and 50 patients in the collected dataset which were manually segmented by expert clinicians.

8.2.1 2DACWE Segmentation Outcome

After the brain tumour location was recognized and identified by BBBGA, the 2DACWE approach was initialized and applied to the T2-w MRI brain slices of a pathological patient from the collected dataset in a slice by slice manner as shown in Fig. 8.2. The ground truth provided by the clinician is marked in green and the tumour boundaries extracted by 2DACWE are marked in red. This patient holds a brain tumour in the right brain hemisphere. The 2DACWE was initialized by the following parameters: $\lambda_1 = \lambda_2 = 1$ and length penalty $\mu = 10^2$. This patient's MRI scan was segmented with a Dice score of 81% by comparing with manual segmentation.

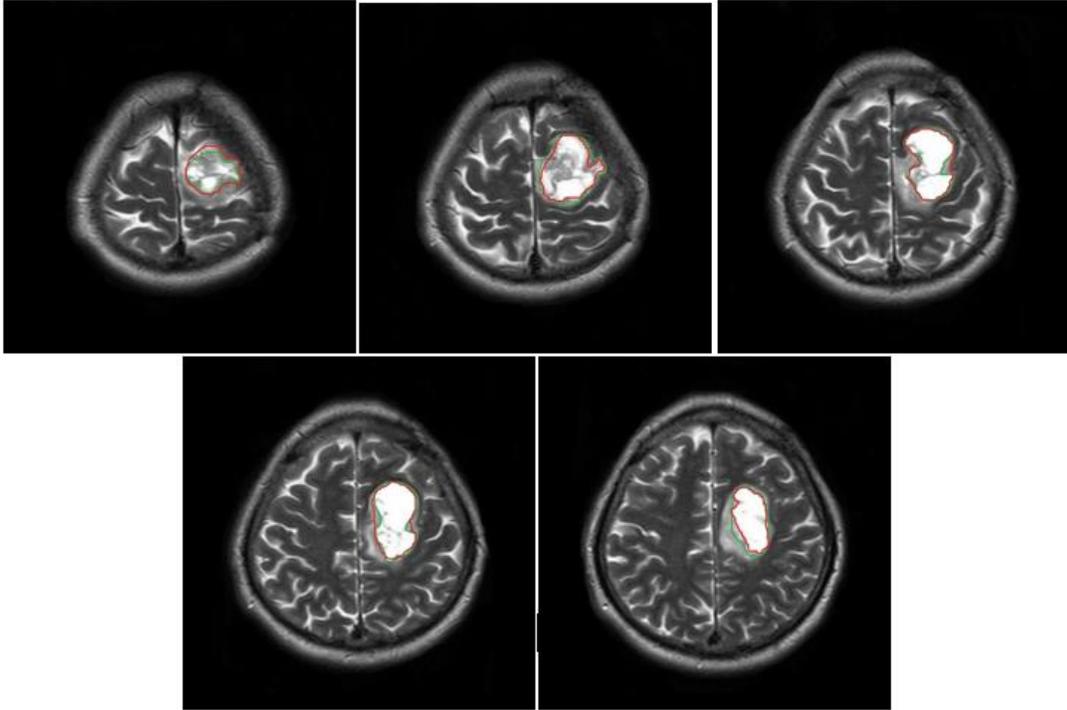


Figure 8.2: Comparative segmentation results on T2-w MRI slices by 2DACWE. The ground truth is marked in green and the output of 3DACWE is marked in red.

8.2.2 3DACWE Segmentation Outcome

After the brain tumour location was recognized and identified by BBBGA, the 3DACWE approach was initialized and applied to 12 T2-w slices of pathological MRI scan from the collected dataset as shown in Fig. 8.3. The ground truth provided by the clinician is marked in green and the tumour boundaries extracted by 3DACWE are marked in red. This patient holds a brain tumour in the left brain hemisphere. The 3DACWE was initialized by the following parameters: $\lambda_1 = \lambda_2 = 1$ and length penalty $\mu = 10^6$. The Chan–Vese energy function was minimized within the iterations of 3DACWE and reached a steady state in 1250 iterations as shown in Fig. 8.4. This patient’s MRI scan was segmented with a Dice score of 88.4% by comparing with manual segmentation.

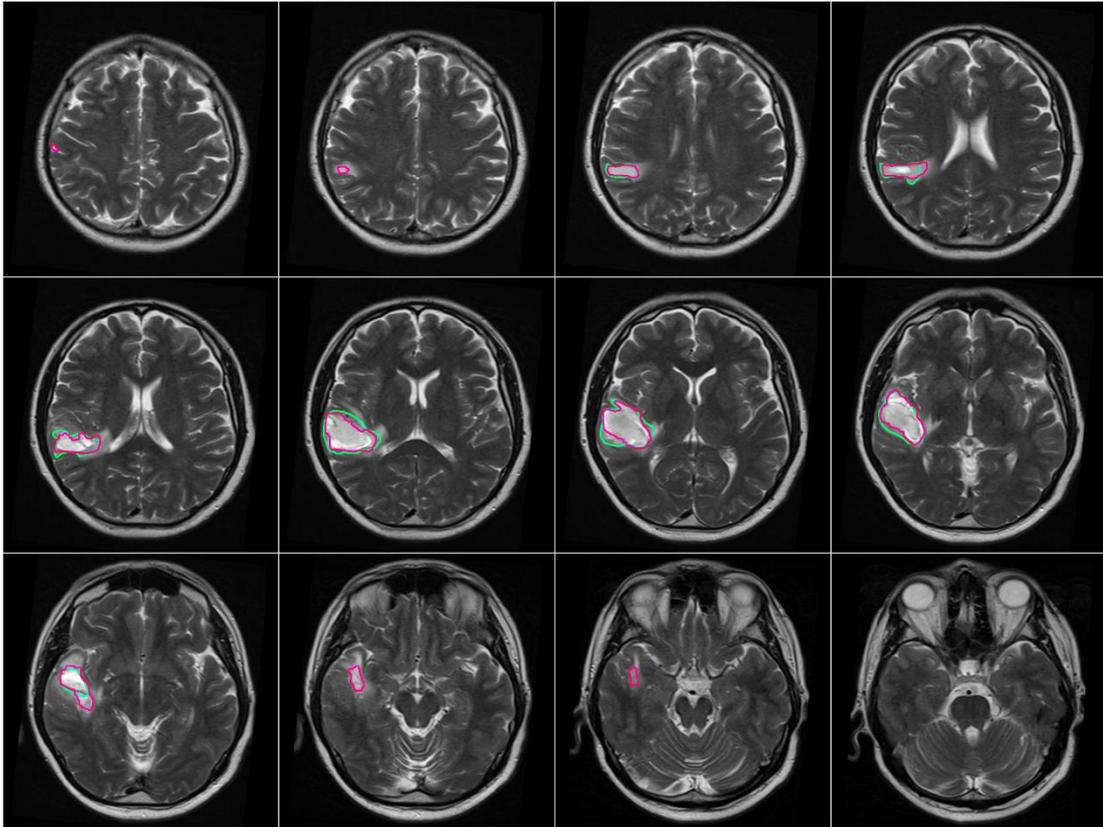


Figure 8.3: Comparative segmentation results on T2-w MRI slices by 3DACWE. The ground truth is marked in green and the output of 3DACWE is marked in red.

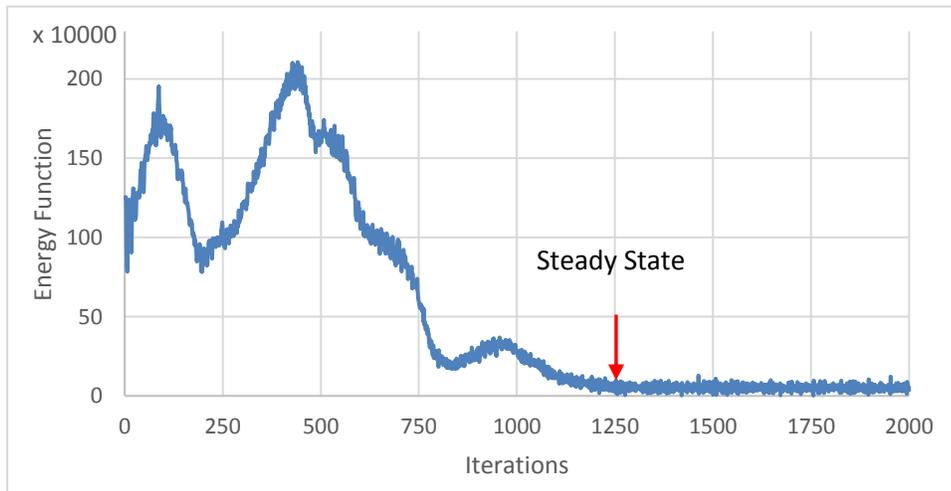


Figure 8.4: The Chan-Vese energy function convergence to steady state.

The 3DACWE approach was initialized and applied to 6 T1-w slices of pathological MRI scan from the collected dataset as shown in Fig. 8.5. The ground truth provided by the clinician is marked in green, and the tumour boundaries extracted by 3DACWE are marked in red. This patient holds a brain tumour in the right brain hemisphere. The 3DACWE was initialized by the following parameters: $\lambda_1 = \lambda_2 = 1$ and length penalty $\mu = 10^6$. The Chan–Vese energy function was minimized within the iterations of 3DACWE and reached a steady state in 1250 iterations as shown in Fig. 8.6. This patient’s MRI scan was segmented with a Dice score of 93.6% by comparing with manual segmentation.

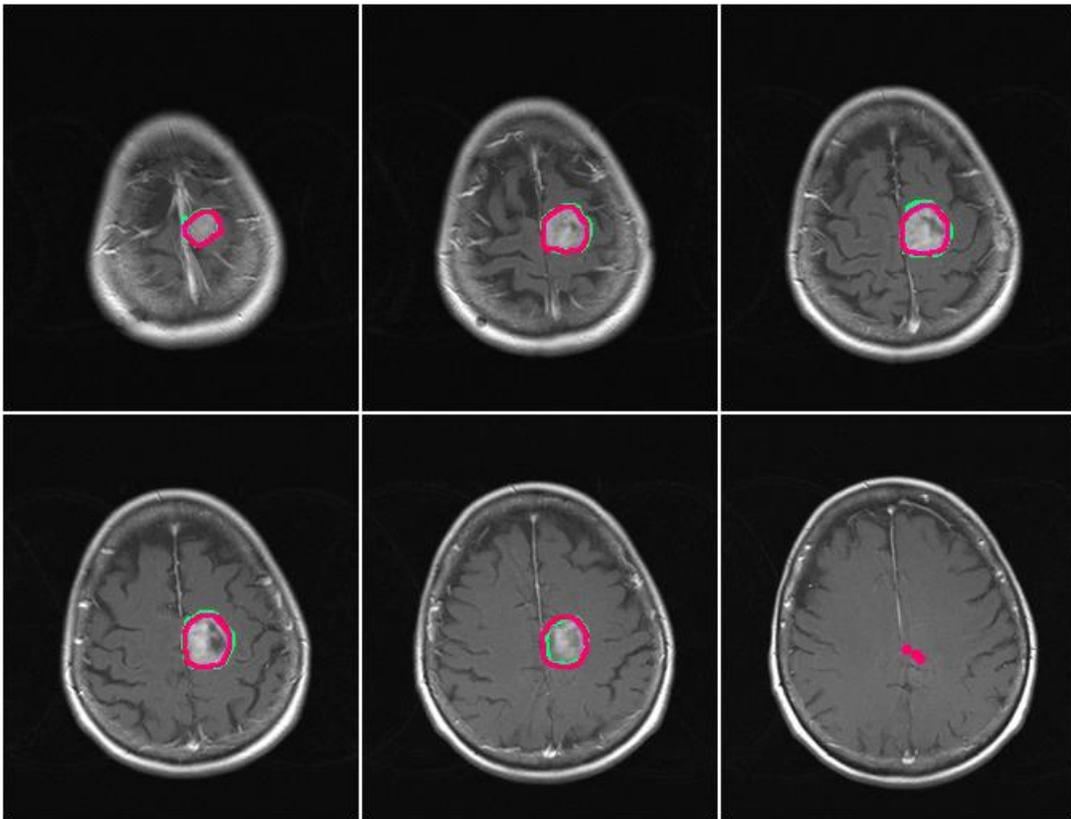


Figure 8.5: Comparative segmentation results on T1-w MRI slices by 3DACWE. The ground truth is marked in green and the output of 3DACWE is marked in red.

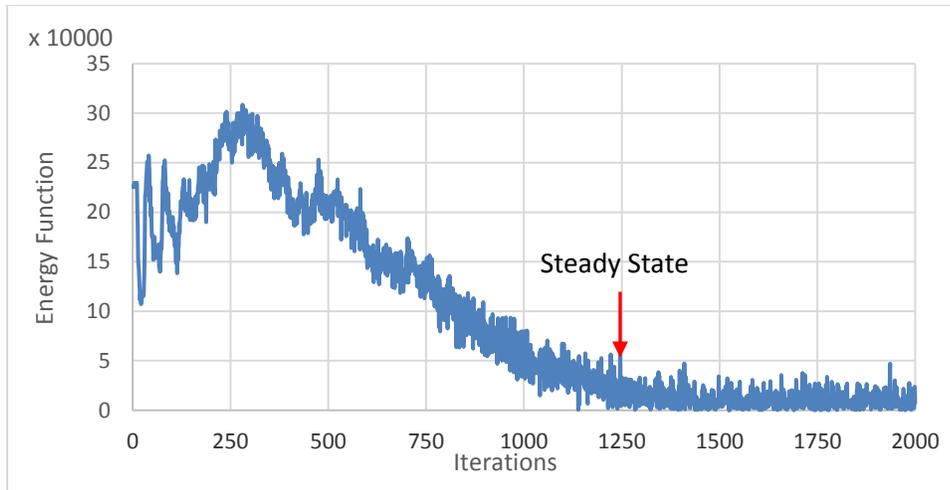


Figure 8.6: The Chan-Vese energy function convergence to steady state.

The 3DACWE approach was initialized and applied to 32 FLAIR slices of pathological MRI scan from the standard dataset (BRATS 2013) as shown in Fig. 8.7. The ground truth provided by BRATS 2013 is marked in green, and the tumour boundaries extracted by 3DACWE are marked in red. This patient holds a brain tumour in the left brain hemisphere. The 3DACWE was initialized by the following parameters: $\lambda_1 = \lambda_2 = 1$ and length penalty $\mu = 10^6$. The Chan-Vese energy function was minimized within the iterations of 3DACWE and reached a steady state in 1750 iterations as shown in Fig. 8.8. This patient's MRI scan was segmented with a Dice score of 94.3% by comparing with the standard dataset of BRATS 2013.

The 3DACWE approach was initialized and applied to 28 T1-w slices of pathological MRI scan from the collected dataset as shown in Fig. 8.9. The ground truth provided by the clinician is marked in green, and the tumour boundaries extracted by 3DACWE are marked in red. This patient holds a brain tumour in the left brain hemisphere. The 3DACWE was initialized by the following parameters: $\lambda_1 = \lambda_2 = 1$ and length penalty $\mu = 10^6$. The Chan-Vese energy function was minimized within the iterations of 3DACWE and reached a steady state in 1100 iterations as shown in Fig. 8.10. This patient's MRI scan was segmented with a Dice score of 93.7% by comparing with manual segmentation.

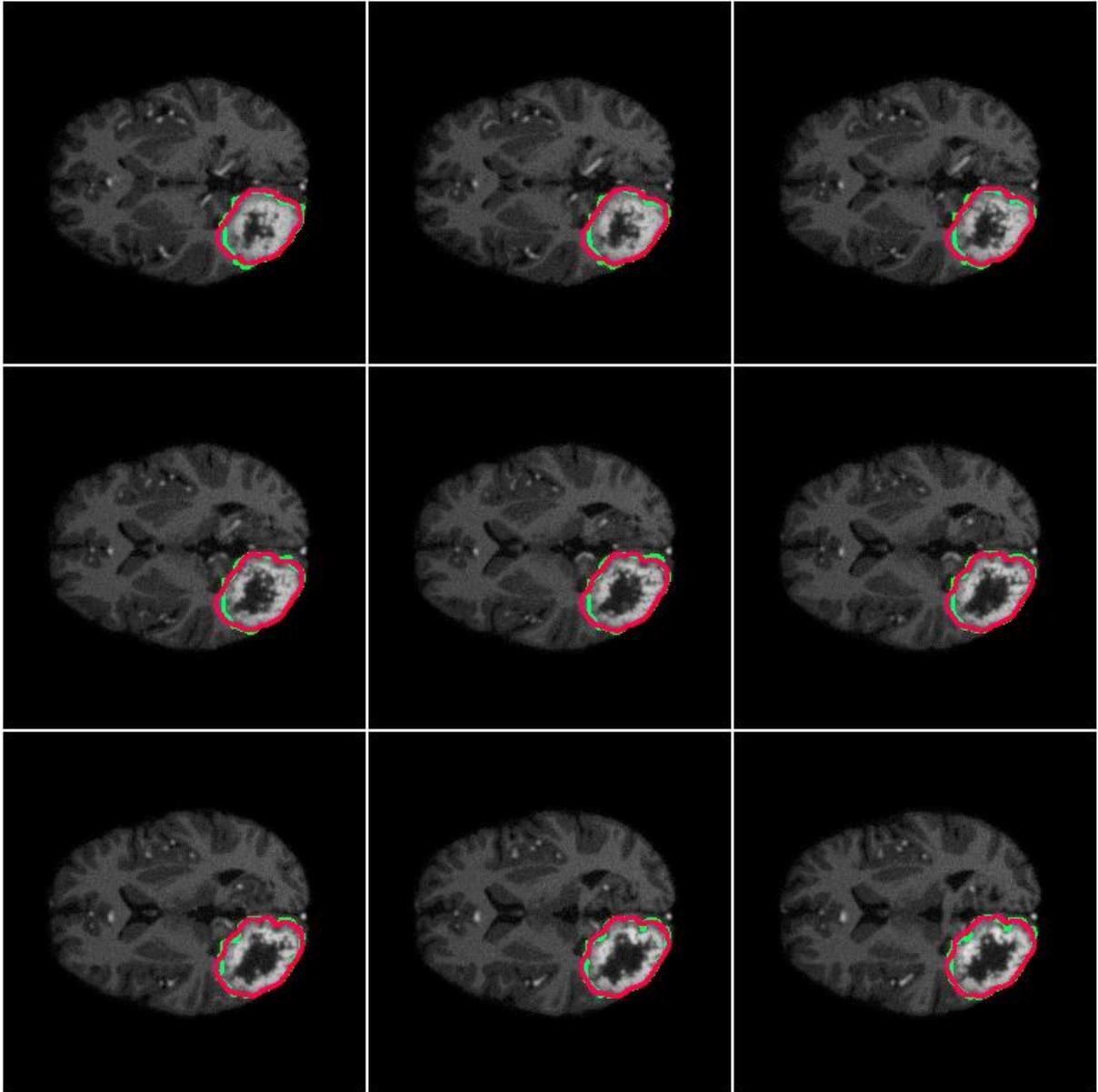


Figure 8.7: Comparative segmentation results on FLAIR slices by 3DACWE. The ground truth is provided by BRATS 2013 in green and the output of 3DACWE is marked in red.

Table 8.1 and Table 8.2 show the results of the segmentation for both collected and BRATS 2013 datasets respectively. The overall results of the segmentation of the four MRI modalities (T1-w, T2-w, T1c-w and FLAIR) for the collected dataset are summarized in Fig. 8.11.

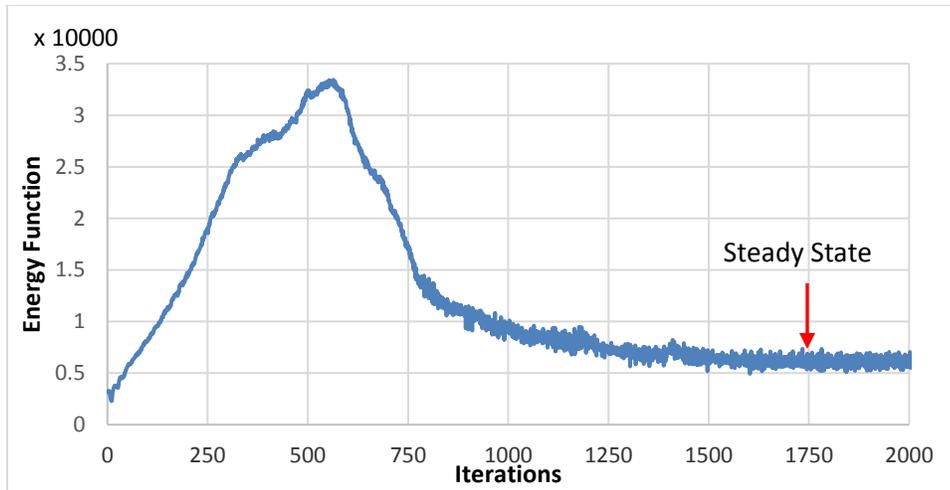


Figure 8.8: The Chan-Vese energy function convergence to steady state.

The T1c-w-based segmentation attained the highest average metric rates because of the contrast enhancement of the pathological tissues. The T2-w-based segmentation was rated as the lowest among all metrics because of highly inhomogeneous intensity distribution despite the sharp edges and high intensity of the brain tumour with respect to the surrounding tissues, where the necrotic tissue of tumour appears with highest intensity and sharp edge. Whilst the edema is of low intensity and with very smooth edges (Mikulka and Gescheidtov, 2013).

Clinically, the delineation of a brain tumour is performed on a single modality of MRI. Lesley Macpherson, a radiologist at the University Hospital Birmingham NHS Foundation Trust gave her opinion about how to delineate brain tumour boundaries in children and concluded that these depends essentially on T2-w, moderately depends on FLAIR and occasionally depends on T1c-w (Tantisatirapong, 2015). While Nigel P. Davies, a lead MRI physicist at the University Hospital Birmingham NHS Foundation Trust, has a different opinion about outlining adult brain tumours. For high grade tumour, he stated that more depend on T1c-w and for outlining low grade tumour, it depends on T2-w and FLAIR (Tantisatirapong, 2015).

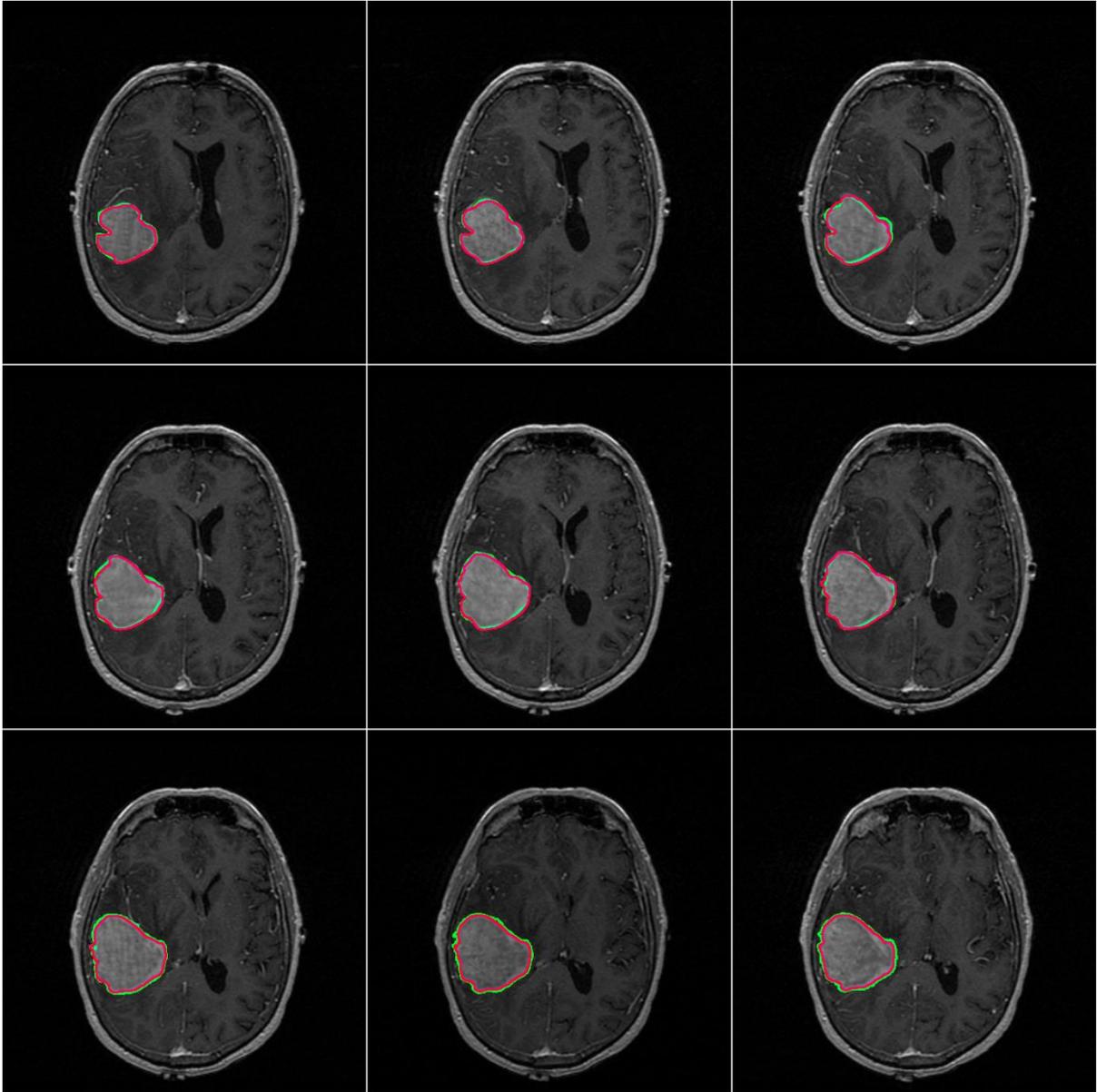


Figure 8.9: Comparative segmentation results on T1-w slices by 3DACWE. The ground truth is marked in green and the output of 3DACWE is marked in red.

Consequently, it is possible to identify the most important slices after implementing 3DACWE method to draw the attention of the clinicians about these slices instead of spending long time on diagnosing and interpreting MRI slices. Table 8.3 shows a comparison of identifying clinically and experimentally the most important slices of 12 pathological patients.

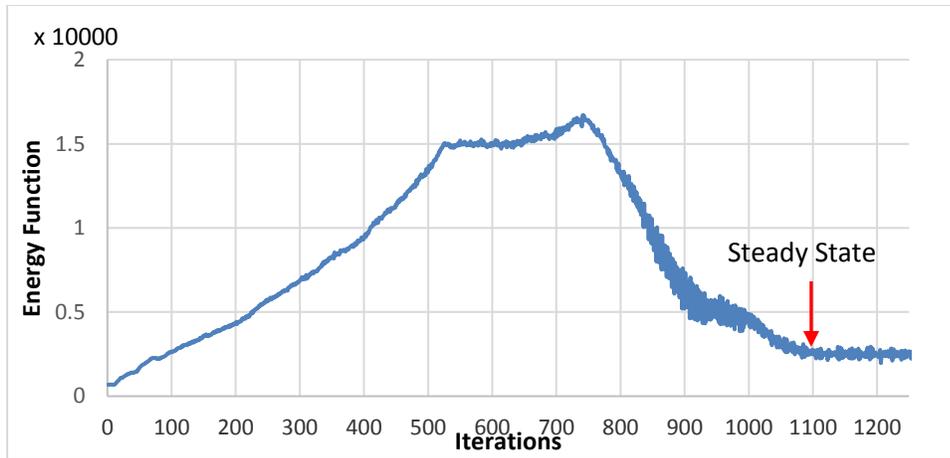


Figure 8.10: The Chan-Vese energy function convergence to steady state.

Table 8.1: Segmentation results for each patient in the collected dataset.

Patient No.	MRI Modality	No. of Slices	TP	FP	TN	FN	Sensitivity	Specificity	Accuracy	Dice Index	Jaccard	Matching
1	T2-w	13	31908	9070	3363603	3291	0.907	0.997	0.996	0.838	0.721	0.907
2	T2-w	12	32699	3562	3104487	4980	0.868	0.999	0.997	0.884	0.793	0.868
3	T2-w	14	84230	6459	3572824	6503	0.928	0.998	0.996	0.929	0.867	0.928
4	T2-w	13	19915	0	3383244	4713	0.809	1.000	0.999	0.894	0.809	0.809
5	T2-w	14	112876	13932	3535931	7277	0.939	0.996	0.994	0.914	0.842	0.939
6	T2-w	11	13767	892	2864378	4547	0.752	1.000	0.998	0.835	0.717	0.752
7	T2-w	15	20511	379	3906820	4450	0.822	1.000	0.999	0.895	0.809	0.822
8	T1-w	8	1862	51	2094601	638	0.745	1.000	1.000	0.844	0.730	0.745
9	T1-w	11	19448	2432	2856952	4752	0.804	0.999	0.998	0.844	0.730	0.804
10	T1c-w	15	16099	417	3911979	3665	0.815	1.000	0.999	0.887	0.798	0.815
11	T1c-w	13	65111	1006	3320724	21031	0.756	1.000	0.994	0.855	0.747	0.756
12	FLAIR	8	18623	3115	2067044	8370	0.690	0.998	0.995	0.764	0.619	0.690
13	T1-w	26	35185	2938	6774629	2992	0.922	1.000	0.999	0.922	0.856	0.922
14	T1-w	34	47567	3860	8853060	8409	0.850	1.000	0.999	0.886	0.795	0.850
15	T2-w	9	5325	1843	2351799	329	0.942	0.999	0.999	0.831	0.710	0.942
16	T1-w	7	2155	229	1831766	858	0.715	1.000	0.999	0.799	0.665	0.715
17	T1-w	13	49476	15051	3341842	1503	0.971	0.996	0.995	0.857	0.749	0.971
18	T2-w	9	15062	3844	2336737	3653	0.805	0.998	0.997	0.801	0.668	0.805
19	T2-w	7	6066	591	1826954	1397	0.813	1.000	0.999	0.859	0.753	0.813
20	T2-w	11	48010	417	2824383	10774	0.817	1.000	0.996	0.896	0.811	0.817
21	T1-w	28	99779	894	7226939	12420	0.889	1.000	0.998	0.937	0.882	0.889
22	T1-w	35	132451	2493	9023249	16847	0.887	1.000	0.998	0.932	0.873	0.887
23	T2-w	7	2007	422	1832022	557	0.783	1.000	0.999	0.804	0.672	0.783
24	T2-w	5	4271	1072	1305133	244	0.946	0.999	0.999	0.867	0.764	0.946
25	T2-w	12	92631	1467	3036437	15193	0.859	1.000	0.995	0.917	0.848	0.859
26	T2-w	12	50081	5571	3080231	9845	0.836	0.998	0.995	0.867	0.765	0.836
27	FLAIR	21	51692	756	5444340	8236	0.863	1.000	0.998	0.920	0.852	0.863
28	T2-w	17	47693	3962	4397953	6840	0.875	0.999	0.998	0.898	0.815	0.875
29	T2-w	18	53821	4944	4656360	3467	0.939	0.999	0.998	0.928	0.865	0.939

30	T1-w	18	28236	1320	4686303	2733	0.912	1.000	0.999	0.933	0.874	0.912
31	T2-w	13	4002	20	3403080	770	0.839	1.000	1.000	0.910	0.835	0.839
32	FLAIR	23	138140	2930	5863542	24700	0.848	1.000	0.995	0.909	0.833	0.848
33	FLAIR	16	144058	2449	4028693	19104	0.883	0.999	0.995	0.930	0.870	0.883
34	T2-w	10	20860	1784	2590147	8649	0.707	0.999	0.996	0.800	0.667	0.707
35	FLAIR	24	195121	1367	6060609	34359	0.850	1.000	0.994	0.916	0.845	0.850
36	T1-w	7	9700	656	1821330	3322	0.745	1.000	0.998	0.830	0.709	0.745
37	T1-w	6	3042	86	1569408	328	0.903	1.000	1.000	0.936	0.880	0.903
38	T1-w	50	66764	1	13031194	9241	0.878	1.000	0.999	0.935	0.878	0.878
39	T2-w	12	11663	498	3130606	2961	0.798	1.000	0.999	0.871	0.771	0.798
40	T1-w	22	8766	13	5756463	1926	0.820	1.000	1.000	0.900	0.819	0.820
41	T1-w	25	120093	794	6420293	12420	0.906	1.000	0.998	0.948	0.901	0.906
42	T2-w	13	80631	867	3311274	15100	0.842	1.000	0.995	0.910	0.835	0.842
43	T1c-w	30	141791	1493	7705236	15800	0.900	1.000	0.998	0.943	0.891	0.900
44	T1c-w	26	40185	938	6770529	4092	0.908	1.000	0.999	0.941	0.889	0.908
45	T2-w	15	60010	1266	3862110	8774	0.872	1.000	0.997	0.923	0.857	0.872
46	T1c-w	17	76001	2179	4365195	13073	0.853	1.000	0.997	0.909	0.833	0.853
47	T1c-w	27	55031	2179	7011605	9073	0.858	1.000	0.998	0.907	0.830	0.858
48	T2-w	18	44925	1337	4666537	5793	0.886	1.000	0.998	0.926	0.863	0.886
49	T2-w	10	14699	962	2605199	580	0.962	1.000	0.999	0.950	0.905	0.962
50	T1c-w	17	20088	1107	2256000	756	0.964	1.000	0.999	0.956	0.915	0.964
Average							0.854	0.999	0.998	0.890	0.804	0.854
STD							0.069	0.001	0.002	0.047	0.075	0.069
Min							0.690	0.996	0.994	0.764	0.619	0.690
Max							0.971	1.000	1.000	0.956	0.915	0.971

Table 8.2: Segmentation results for each patient in the BRATS 2013 dataset.

Patient No.	MRI Modality	No. of Slices	TP	FP	TN	FN	Sensitivity	Specificity	Accuracy	Dice Index	Jaccard	Matching
1	T1c-w	23	3992	8	6024139	1173	0.773	1.000	1.000	0.871	0.772	0.773
2	FLAIR	42	19832	1075	10987391	1750	0.919	1.000	1.000	0.934	0.875	0.919
3	T1c-w	26	4698	488	6810273	285	0.943	1.000	1.000	0.924	0.859	0.943
4	FLAIR	53	55793	1065	13819891	16883	0.768	1.000	0.999	0.861	0.757	0.768
5	T2-w	48	26109	1115	12546345	9343	0.736	1.000	0.999	0.833	0.714	0.736
6	T2-w	24	7752	3131	6279025	1548	0.834	1.000	0.999	0.768	0.624	0.834
7	FLAIR	49	50448	3490	12788270	2848	0.947	1.000	1.000	0.941	0.888	0.947
8	T2-w	56	49417	730	14617608	12309	0.801	1.000	0.999	0.883	0.791	0.801
9	T1c-w	72	236141	16438	18611652	10137	0.959	0.999	0.999	0.947	0.899	0.959
10	T2-w	49	44418	6014	12792678	1946	0.958	1.000	0.999	0.918	0.848	0.958
11	T2-w	43	34509	7401	11229012	1270	0.965	0.999	0.999	0.888	0.799	0.965
12	T1c-w	58	101547	5968	15090734	6103	0.943	1.000	0.999	0.944	0.894	0.943
13	T1c-w	45	34156	1695	11757396	3233	0.914	1.000	1.000	0.933	0.874	0.914
14	T1c-w	43	39138	8707	11224152	195	0.995	0.999	0.999	0.898	0.815	0.995
15	T2-w	49	49259	11777	12777738	6282	0.887	0.999	0.999	0.845	0.732	0.887
16	FLAIR	46	53545	3826	11997820	3433	0.940	1.000	0.999	0.937	0.881	0.940
17	T1c-w	36	23651	1243	9409786	2504	0.904	1.000	1.000	0.927	0.863	0.904
18	T1c-w	37	26750	5782	9666577	219	0.992	0.999	0.999	0.899	0.817	0.992
19	T1c-w	41	34498	5694	10706266	1446	0.960	0.999	0.999	0.906	0.829	0.960
20	T1c-w	39	25586	5480	10191639	911	0.966	0.999	0.999	0.889	0.800	0.966

21	T1c-w	40	19980	4654	10458884	2242	0.899	1.000	0.999	0.853	0.743	0.899
22	T1c-w	34	21985	4500	8886231	180	0.992	0.999	0.999	0.904	0.824	0.992
23	T1c-w	33	19156	3576	8626033	1987	0.906	1.000	0.999	0.873	0.775	0.906
24	T1c-w	41	28023	0	10714312	5569	0.834	1.000	0.999	0.910	0.834	0.834
25	FLAIR	56	64018	24354	14591612	80	0.999	0.998	0.998	0.840	0.724	0.999
Average							0.909	1.000	0.999	0.893	0.809	0.909
STD							0.076	0.000	0.000	0.043	0.068	0.076
Min							0.736	0.998	0.998	0.768	0.624	0.736
Max							0.999	1.000	1.000	0.947	0.899	0.999

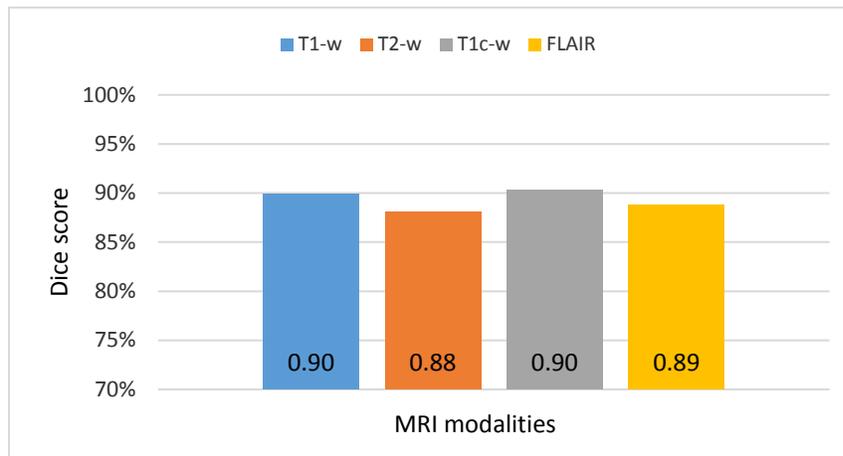


Figure 8.11: The overall results of the segmentation of the four MRI modalities.

Table 8.3: A comparison between clinical and experimental slice identification.

Patient No.	Clinically Identified		Experimentally Identified	
	Tumour Starting Slice	Tumour Ending Slice	Tumour Starting Slice	Tumour Ending Slice
1	5	11	4	11
2	4	11	4	11
3	3	11	3	11
4	3	9	3	9
5	2	12	3	11
6	5	14	5	14
7	8	13	8	13
8	16	22	16	22
9	6	10	6	10
10	4	10	5	10
11	2	11	2	11
12	1	13	1	13

In this study, the three-dimensional segmentation has been proved that is more successful than the two-dimensional segmentation due to the greater number of pixels that occupied a volume inside and outside the boundary of brain tumours that are used within the segmentation process. The achieved average Dice score of the 2DACWE method (slice by slice) on the collected dataset was $84.18 \pm 4.5\%$. Fig. 8.12 shows a comparison of average Dice scores of four modalities of MRI using 2DACWE and 3DACWE segmentation techniques.

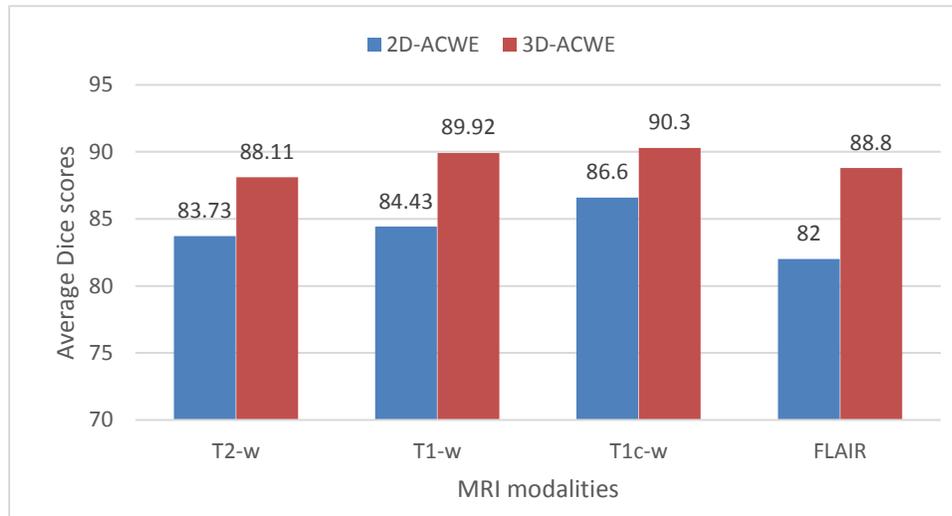


Figure 8.12: Comparing tumour segmentation average Dice scores of four modalities of MRI using 2DACWE and 3DACWE.

Figure 8.13 illustrates the comparison between the 3DACWE and 2DACWE segmentation results of the collected dataset of 50 pathological patients. Notably, 3DACWE outweighs the 2DACWE method for all patients in the given dataset due to exploiting all important anatomical information and the features of the full MRI scans.

Tables 8.4 show a comparison between the existing segmentation methods and our proposed segmentation method based on number of patients (training/testing), Accuracy, Match, Jaccard and Dice evaluators in different MRI modalities. It is observed from the table that the best performance is for proposed segmentation method with Dice score of 89 % for 50 patients of the collected dataset and 89.3% for BRATS 2013 dataset. Generally, it is observed from the comparison table the 3DACWE proposed method has achieved the best performance in low computational time with high segmentation Dice score compared to the other recently developed MRI image segmentation techniques.

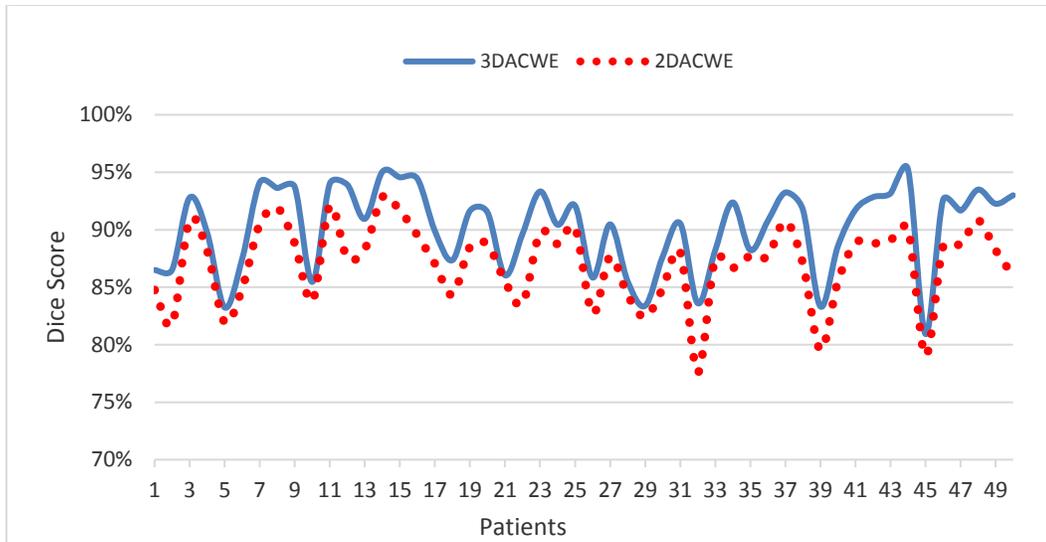


Figure 8.13: Comparison between 3DACWE and 2DACWE segmentation results for collected dataset (50 patients).

8.3 Conclusion

The quantitative measures of the automated 3DACWE segmentation come closer to the manual expert segmentation. Such that, the achieved average Dice score of segmenting the collected dataset was $89 \pm 4.7\%$ with a sensitivity rate of 85.4%. While the achieved average Dice score of BRATS 2013 dataset was $89.3 \pm 4.3\%$ with a sensitivity rate of 91%. It was noted that the segmentation accuracy of 3DACWE decreases significantly with increasing in the summation of slice thickness and space between slices as shown in Fig. 8.14. The scatter plot shows a negative correlation between the Dice score and the summation of slice thickness and space between slices. Therefore, to achieve a high segmentation accuracy, a reduction of the slice thickness and space between slices to a minimum is essential and diminishes the PV effect.

We conclude that the 3DACWE method is effective in brain tumour segmentation because the approach does not only consider local tumour properties (gradients), but also relies on global properties (intensity), contour length and region length. Although the achieved accuracy was high relative to state-of-arts segmentation techniques, the 3DACWE was relatively slow for brain tumour segmentation due to processing of a volumetric data of MRI of 512×512 pixel resolution.

Table 8.4: An overview of segmentation methods compared to proposed system.

Reference	MRI Modalities	Approach	No. of Patients (training/Testing)	Performance			
				Accuracy	Match	Jaccard	Dice
(Fletcher-Heath et al., 2001)	T1, T2 and PD	Fuzzy clustering	2/4	-	53-91%	-	-
(Kaus et al., 2001)	T1	Template-moderated classification	10/10	95%	-	-	-
(Ho et al., 2002)	T1 and T1-c	Level-sets	Na./5	-	-	85-93%	-
(Prastawa et al., 2004)	T2	Generative model	Na./3	-	-	59-89%	-
(Corso et al., 2008)	T1, T1c, T2 and FLAIR	Weighted aggregation	10/10	-	-	62-69%	-
(Verma et al., 2008)	T1, T1c, T2 and FLAIR	SVM	14	34-93%	-	-	-
(Menze et al., 2010)	T1, T1c, T2 and FLAIR	Generative model	25	-	-	-	40-70%
(Havaei et al., 2017)	T1, T1c, T2 and FLAIR	Deep Neural Network	30/10 (BRATS 2013)	-	-	-	88%
(Nabizadeh, 2015)	T1, T1c, T2 and FLAIR	Skippy greedy snake	25 (BRATS 2013)	96.8±0.3%	-	-	-
(Guo et al., 2015)	T1	SVM	60	98.5±1.1%	-	-	73.1±10.6%
			30 (BRATS 2012)	99.2±1.1%			66.5±12%
Proposed System	T1, T1c, T2 and FLAIR	3DACWE	50 (collected dataset)	99.8±0.2%	85.4±6.9%	80.4±7.4%	89±4.7%
			25 (BRATS 2013)	99.9%	91±7.6%	81±6.8%	89.3±4.2%

A major difficulty was encountered during white matter tumours segmentation because of the overlapping WM/GM intensity distributions in such cases. Some parts of the tumours in GM were not distinguished because of the restricted image resolution and the complex network of brain tissue with various shapes and sizes. These factors significantly affect many voxels located in the tissue borders. Moreover, central tumour image intensity slightly differs from the peripheral tumour intensity. As such, the intensity near the borders can be considered similar to that of the grey matter. Consequently, the tumour and GM may be confused and the peripheral tumour regions may be misclassified.

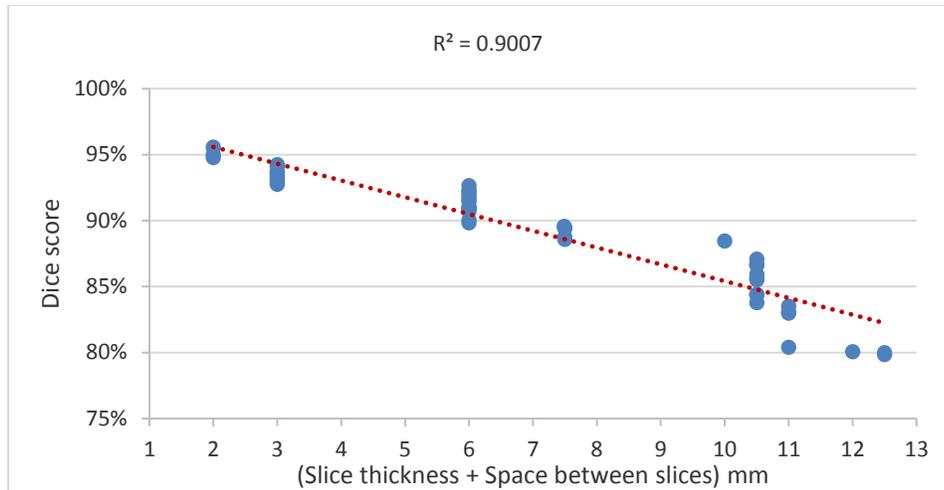


Figure 8.14: Scatter plot of Dice score to the summation of slice thickness and space between slices, showing mean of Dice score (R^2) as the dotted red line.

Tumour size affects segmentation accuracy and errors usually occur at the tumour boundaries. Large brain tumours contain a high number of image pixels that can be misclassified. Moreover, large tumours likely ingress into the brain boundary and CSF and render the precise determination of the tumour boundary challenging. With regard to the overall execution time, the proposed system handled volumetric MRI data with different characteristics, such as the number of slices and tumour size, type, boundary and location. These characteristics make the overall segmentation process time consuming. Hence, the processing time of the proposed system was measured by second per MRI slice. An average processing time of our proposed system requires 35 s/MRI slice to run the segmentation.

CHAPTER NINE

Conclusion and Future Work

Overview

This chapter concludes and summarizes the work that has been developed as part of the PhD programme and the techniques that have been investigated to discriminate and differentiate pathological MRI brain scans from the normal brain scans and segment the brain tumours. The main achievements are discussed and suggestions for future work are made.

9.1 Introduction

Medical image processing has expanded dramatically during the last decade and has been an interesting research field that attracted expertise from other fields such as mathematics, computer sciences, engineering, biology and medicine. The main aim of this study is to come out with an automated system for the classification and segmentation of brain tumours in MRI scans. These aspects were considered and addressed in this study that has scientifically contributed in many ways. Most of previous studies in the brain tumour segmentation field were based on multi-spectral MRI, multi-scale classification, local or global registration, high resolution and non-noisy data. The research was conducted through its objectives as presented in section 1.3.

The first and second objectives were developed as described in chapters 2 and 4 respectively, through evaluating and comparing recent algorithms of brain tumours classification and segmentation of MRI brain scans. In chapter 2, several related works and methods were presented, investigated, analysed, and the reader was given a comprehensive knowledge about different related techniques, in addition to advantages and disadvantages. While, chapter 4 demonstrates how to prepare the MRI brain slices to achieve high brain tumours classification and segmentation accuracies, a set of image pre-processing algorithms were implemented on the collected dataset from Al-Kadhimiya Teaching Hospital in Iraq. These algorithms were resizing the dimensions of MRI slices, MRI image enhancement, intensity normalization,

background segmentation and MSP detection and correction algorithms. The third objective was also achieved as described in chapters 5 and 6. This objective represents the most important part of this research that aimed to design and develop an efficient automated screening system of MRI brain scanning. Two novel methods were proposed to extract the second order statistical features that were used to discriminate and classify the MRI brain scans into normal and pathological brain scanning in high enough accuracy and with a reasonable computation complexity. The first method was called modified grey level co-occurrence matrix (MGLCM), which was based on using single modality of MRI (T2-w images) and exploited the symmetry property of the two hemispheres of a normal brain. Nine matrices and 100 co-occurrence statistics were extracted from each MRI slice after eliminating the irrelevant and redundant features by ANOVA. The remarkable accuracy of this method in classifying the MRI brain scans by MLP ($97.8 \pm 0.1\%$ for the collected dataset and $98.6 \pm 0.15\%$ for the standard dataset) demonstrates its efficiency. The second method was called three-dimensional modified grey level co-occurrence matrix (3DMGLCM) and was based on using a single volumetric modality of MRI (T2-w images). It has the same principles of MGLCM method but the difference is that the 3DMGLCM works on all MRI brain slices at the same time, while the MGLCM works on each MRI slice separately. Nine matrices and 100 co-occurrence statistics were extracted from volumetric data of MRI scan after eliminating the irrelevant and redundant features by ANOVA. The accuracy of this method in classifying the volumetric data of MRI by MLP ($93.3 \pm 0.15\%$ for the collected dataset and $95.3 \pm 0.2\%$ for the standard dataset) was very high.

The fourth objective was developed in chapter 7 where a new method was proposed and named bounding 3D-boxes based genetic algorithm (BBBGA) for localizing the brain tumours of pathological patients by generating a hundred of 3D-boxes within the volumetric data of MRI scan. These 3D-boxes are being moved during the iterations of GA toward the area that has less similarity representing a candidate tumour region. Then the final location of 3D-boxes can be used to construct the initial seed points inside the brain tumour to initialize the segmentation algorithm automatically without the need for user interaction as presented in chapter 8, which represents the fifth objective. The main advantages of BBBGA is being an unsupervised method and there is no need for atlas registration, intensity standardization of MRI slices and skull removing as well as it uses a single modality of MRI (T2-w images). The achieved accuracy by BBBGA was 95%, as there were only 4 cases whose tumour's size is less than 1 cm^3 , have failed

to identify the abnormality because of the relative similarity of spatial scale between the tiny brain tumours and normal variability in anatomy. Chapter 8 included the implementation of the 2DACWE and 3DACWE respectively. A comparison was made between these two methods of segmentation and the best segmentation accuracy of brain tumours was achieved by 3DACWE because it exploits the entire volumetric MRI data within the segmentation process instead of segmenting each slice separately. The achieved average Dice scores for segmenting the collected dataset and the standard dataset (BRATS 2013) were $89\pm 4.7\%$ and $89.3\pm 4.3\%$ respectively. Furthermore, a comparison was made with other state-of-the-art of classification and segmentation of brain tumours in MRI scans. This comparison stated that the proposed system outperformed the other recently developed systems.

The overall automated system in this study has some advantages that can be summarised as follows:

- It is independent of atlas registration, as depending on atlas registration means that the performance depends on how well the atlas is constructed and how well the registration algorithm can register the test data to the atlas.
- It is independent of prior anatomical knowledge, as the dependence means that the algorithm must be trained to incorporate such information which can lead to errors in classification and segmentation.
- The proposed system does not need to initialize assumptions about the number of classes in MRI scan.
- All the proposed methods are fully automatic.

9.2 Future Works

According to the conclusion and contribution that are provided in this study, there is a set of recommendations for the researchers who are interested in brain tumours classification and segmentation. The future developments of this work can be summarised as follows:

- Further studies should involve in the pre-processing stage that can be improved and enhanced the performance of the classification and segmentation of the system.

- Investigate and quantify the features that are extracted from other modalities of MRI such as T1-w, T1c-w and FLAIR slices and combine them with the features that were employed in this study to classify the pathological MRI scans further according to the various types of brain tumours, such as primary gliomas from metastases, grading of gliomas, very small tumours, lesions caused by injury and dementia.
- Reduction of computer processing time is another important issue, because the clinical diagnostic routine applications should not exceed the order of minutes. Therefore, it is essential to be optimized with respect to faster implementation in computers network that will run in parallel mode.

REFERENCES

- Aelterman, J., Goossens, B., Pizurica, A. & Philips, W. Removal of correlated rician noise in magnetic resonance imaging. 16th European Signal Processing Conference, 2008, Switzerland. pp. 1-5.
- Agrawal, R. & Sharma, M. 2014. Review of segmentation methods for brain tissue with magnetic resonance images. International Journal of Computer Network and Information Security, vol. 6, no. 4, pp. 55-62.
- Ain, Q., Mehmood, I., Naqi, S. M. & Jaffar, M. A. Bayesian classification using DCT features for brain tumor detection. Proceedings of the Knowledge-Based and Intelligent Information and Engineering Systems (KES), 2010 UK. Springer-Verlag Berlin Heidelberg, pp. 340-349.
- Al Hilfi, T. K., Lafta, R. & Burnham, G. 2013. Health services in Iraq. The Lancet, vol. 381, no. 9870, pp. 939-948.
- Albregtsen, F. 2008. Statistical Texture Measures Computed from Gray Level Cooccurrence Matrices. Image processing laboratory, department of informatics, university of oslo, vol. 5, no. 2008, pp. 1-14.
- Aloui, K. & Naceur, M. S. 2011. 3D tumor segmentation from volumetric brain MR images using level-sets method. ANA, L. (ed.). InTech, DOI: 10.5772/19436.
- Alwan, A. 2004. Health in Iraq. Minister of Health-Iraq. Available: http://www.who.int/hac/crises/irq/background/Iraq_Health_in_Iraq_second_edition.pdf?ua=1
- American Brain Tumor Association. 2015. SURGERY. USA. Available: <http://www.abta.org/secure/surgery.pdf>
- Anbeek, P., Vincken, L., Van Bochove, S., Van Osch, P. & Van Der Grond, J. 2005. Probabilistic segmentation of brain tissue in MR imaging. NeuroImage, vol. 27, no. 4, pp. 795-804.
- Anju, E., Karnan, M. & Sivakumar, R. 2013. MR brain image classification with supervised bacteria foraging technique using SVM. International Journal of Futuristic Science Engineering and Technology, vol. 1, no. 5, pp. 318-322, .
- Antkowiak, M. 2006. Artificial Neural Networks vs. Support Vector machines for skin diseases recognition. Master Degree, Department of Computing Science, Umea University, Sweden, vol., no., pp.
- Ardekani, A., Kershaw, J., Braun, M. & Kanno, I. 1997. Automatic detection of the mid-sagittal plane in 3-D brain images. IEEE transactions on medical imaging, vol. 16, no. 6, pp. 947-52.
- Ashton, E. A., Parker, K. J., Berg, M. J. & Chen, C. W. 1997. A novel volumetric feature extraction technique with applications to MR images. IEEE Transactions on Medical Imaging, vol. 16, no. 4, pp. 365-371.
- Baaziz, N., Abahmane, O. & Missaoui, R. 2010. Texture feature extraction in the spatial-frequency domain for content-based image retrieval. arXiv preprint arXiv:1012.5208, vol., no., pp.
- Babatunde, O., Armstrong, L., Leng, J. & Diepeveen, D. 2014. A genetic algorithm-based feature selection. British Journal of Mathematics & Computer Science, vol. 4, no. 21, pp. 889-905.

- Baboo, S. & Sasikala, S. 2010. Multicategory Classification Using Support Vector Machine for Microarray Gene Expression Cancer Diagnosis. *Global Journal of Computer Science and Technology*, vol. 10, no. 15, pp. 38-44.
- Balakrishnama, S. & Ganapathiraju, A. 1998. Linear discriminant analysis-a brief tutorial. Institute for Signal and information Processing, Mississippi State University, USA. Available: [https://datajobs.com/data-science-repo/LDA-Primer-\[Balakrishnama-and-Ganapathiraju\].pdf](https://datajobs.com/data-science-repo/LDA-Primer-[Balakrishnama-and-Ganapathiraju].pdf)
- Bandyopadhyay, S. & Pal, S. 2007. *Classification and Learning Using Genetic Algorithms: Applications in Bioinformatics and Web Intelligence (Natural Computing Series)*, Springer-Verlag New York, Inc.
- Bankman, I. N. 2009. Segmentation. *Handbook of Medical Image Processing and Analysis (Second Edition)*. Burlington: Academic Press.
- Bankman, I. N., Spisz, T. S. & Pavlopoulos, S. 2009. Chapter 15 - Two-Dimensional Shape and Texture Quantification. *Handbook of Medical Image Processing and Analysis (Second Edition)*. Burlington: Academic Press.
- Bauer, S., Nolte, L. & Reyes, M. Fully Automatic Segmentation of Brain Tumor Images using Support Vector Machine Classification in Combination with hierarchical Conditional Random Field Regularization. *Medical Image Computing and Computer-Assisted Intervention, 2011*, Springer Berlin Heidelberg. pp. 354-361.
- Belkic, D. & Belkic, K. 2010. *Signal Processing in Magnetic Resonance Spectroscopy with Biomedical Applications*, CRC Press.
- Beneš, M. & Zitová, B. 2015. Performance evaluation of image segmentation algorithms on microscopic image data. *Journal of microscopy*, vol. 257, no. 1, pp. 65-85.
- Bergo, F., Falcão, A., Yasuda, C. & Ruppert, G. Fast, Accurate and Precise Mid-Sagittal Plane Location in 3D MR Images of the Brain. In *International Joint Conference on Biomedical Engineering Systems and Technologies, 2008*. Springer Berlin Heidelberg, pp. 278-290.
- Berry, E. 2007. *A practical approach to medical image processing*, Taylor & Francis.
- Beura, S., Majhi, B. & Dash, R. 2015. Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer. *Neurocomputing*, vol. 154, no., pp. 1-14.
- Birry, R. 2013. *Automated Classification in Digital Images of Osteogenic differentiated stem Cells*. PhD, University of Salford, Manchester, UK.
- Blink, E. 2004. *Basic mri : Physics*.
- Bovik, A. 2009. *The Essential Guide to Image Processing*, Elsevier Inc.
- Brody, W. R. 2009. Foreword A2 - BANKMAN, ISAAC N. *Handbook of Medical Image Processing and Analysis (Second Edition)*. Burlington: Academic Press.
- Burns, R. & Burns, R. 2008. *Business Research Methods and Statistics Using SPSS*, SAGE Publications Ltd.
- Chan, F., Sandberg, B. & Vese, A. 2000. Active contours without edges for vector-valued images. *Journal of Visual Communication and Image Representation*, vol. 11, no. 2, pp. 130-141.
- Chan, F. & Vese, A. 2001. Active contours without edges. *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266-277.
- Chen, D., Liu, Z., Ma, X. & Hua, D. 2005. Selecting genes by test statistics. *BioMed Research International*, vol. 2005, no. 2, pp. 132-138.

- Chen, S. & Tian, Y. Evaluating effectiveness of Latent Dirichlet Allocation model for scene classification. Wireless and Optical Communications Conference (WOCC), 2011, USA. IEEE, pp. 1-6.
- Chipperfield, A., Fleming, P., Pohlheim, H. & Fonseca, C. 1994. Genetic Algorithm TOOLBOX For Use with MATLAB, UK, University of Sheffield.
- Christensen, J., Hutchins, G. & McDonald, C. 2006. Computer automated detection of head orientation for prevention of wrong-side treatment errors. American Medical Informatics Association Annual Symposium Proceedings,, vol. 2006, no., pp. 136.
- Christos, L. 2005. Ultrasound image analysis of the carotid artery. PhD, Kingston University, London, UK.
- Chunming, L., Chiu-Yen, K., Gore, C. & Zhaohua, D. 2008. Minimization of Region-Scalable Fitting Energy for Image Segmentation. IEEE Transactions on Image Processing, vol. 17, no. 10, pp. 1940-1949.
- Coley, D. A. 1999. An introduction to genetic algorithms for scientists and engineers, World scientific.
- Colin, R. & Jonathan, E. 2002. Genetic algorithms-Principles and perspectives, A guide to GA Theory, USA, Kluwer Academic Publisher.
- Corso, J. J., Sharon, E., Dube, S., El-Saden, S., Sinha, U. & Yuille, A. 2008. Efficient multilevel brain tumor segmentation with integrated bayesian model classification. IEEE Transactions on Medical Imaging, vol. 27, no. 5, pp. 629-640.
- Crandall, R. 2009. Image segmentation using the Chan-Vese algorithm. Project report from ECE. Available: http://math.arizona.edu/~rcrandall/ECE532_ProjectPaper.pdf
- Crawley, M. J. 2012. The R book, John Wiley & Sons.
- Despotović, I., Goossens, B. & Philips, W. 2015. MRI segmentation of the human brain: challenges, methods and applications. Computational and mathematical methods in medicine, vol. 2015, no., pp. 23.
- Dou, W., Ruan, S., Chen, Y., Bloyet, D. & Constans, J.-M. 2007. A framework of fuzzy information fusion for the segmentation of brain tumor tissues on MR images. Image and Vision Computing, vol. 25, no. 2, pp. 164-171.
- Dougherty, G. 2009. Digital Image Processing for Medical Applications, Cambridge University Press.
- Drevelegas, A. & Papanikolaou, N. 2011. Imaging modalities in brain tumors. ANTONIOS, D. (ed.) Imaging of brain tumors with histological correlations. Springer Berlin Heidelberg.
- Du, C. J., Iqbal, A. & Sun, D. W. 2016. Chapter 8 - Quality Measurement of Cooked Meats. SUN, D.-W. (ed.) Computer Vision Technology for Food Quality Evaluation (Second Edition). San Diego: Academic Press.
- Dubitzky, W., Granzow, M. & Berrar, D. P. 2007. Fundamentals of data mining in genomics and proteomics, Springer Science & Business Media.
- Dvořák, P., Kropatsch, W. & Bartušek, K. 2013. Automatic brain tumor detection in t2-weighted magnetic resonance images. The Journal of Institute of Measurement Science of Slovak Academy of Sciences, vol. 13, no. 5, pp. 223-230.
- El-Shenawy, M. 2013. Automatic detection and identification of cells in digital images of day 2 IVF embryos. PhD, University of Salford, Manchester, UK.
- Eleyan, A. & Demirel, H. Co-occurrence based statistical approach for face recognition. 24th International Symposium on Computer and Information Sciences (ISCIS) 2009. IEEE, pp. 611-615.

- Fabijańska, A. Results of applying two-pass region growing algorithm for airway tree segmentation to MDCT chest scans from EXACT database. The Second International Workshop on Pulmonary Image Analysis, CreateSpace, USA, 2009. pp. 251-260.
- Fathi, R., Matti, L., Al-Salih, H. & Godbold, D. 2013. Environmental pollution by depleted uranium in Iraq with special reference to Mosul and possible effects on cancer and birth defect rates. *Medicine, Conflict and Survival*, vol. 29, no. 1, pp. 7-25.
- Fletcher-Heath, L. M., Hall, L. O., Goldgof, D. B. & Murtagh, F. R. 2001. Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. *Artificial intelligence in medicine*, vol. 21, no. 1, pp. 43-63.
- Gadkari, D. 2004. Image quality analysis using GLCM. M.Sc, University of Central Florida, USA.
- Gebejes, A. & Huertas, R. Texture Characterization based on Grey-Level Co-occurrence Matrix. In *Proceeding of Conference of Informatics and Management Sciences*, 2013 Slovakia. EDIS - Publishing Institution of the University of Zilina, pp. 375-378.
- Getreuer, P. 2012. Chan-vede segmentation. *Image Processing On Line*, vol. 2, no., pp. 214-224.
- Gomez, W., Pereira, W. & Infantosi, A. 2012. Analysis of Co-Occurrence Texture Statistics as a Function of Gray-Level Quantization for Classifying Breast Ultrasound. *IEEE Transactions on Medical Imaging*, vol. 31, no. 10, pp. 1889-1899.
- Gonzalez, R. & Woods, R. 2002. *Digital Image Processing*, Prentice-Hall, Inc.
- Gordillo, N., Montseny, E. & Sobrevilla, P. 2013. State of the art survey on MRI brain tumor segmentation. *Magnetic Resonance Imaging*, vol. 31, no. 8, pp. 1426-1438.
- Graupe, D. 2013. *Principles of Artificial Neural Networks*, World Scientific.
- Grefenstette, J. 1986. Optimization of Control Parameters for Genetic Algorithms. *IEEE Transactions on systems, man and cybernetics*, vol. 16, no. 1, pp. 122-128.
- Gunasekara, R., Geethal, B., Hassan Mafaz, M., Mathew, C., Perera, A., Subasinghe, H., Welikala, Y. & Wimalasena, L. 2009. Prophetia: Artificial Intelligence for TravelBox® Technology. W. YU & E. N. SANCHEZ (eds.) *Advances in Computational Intelligence*. Springer-Verlag Berlin Heidelberg.
- Günther, F. & Fritsch, S. 2010. neuralnet: Training of Neural Networks. *The R Journal*, vol. 2, no. 1, pp. 30-38.
- Guo, D., Fridriksson, J., Fillmore, P., Rorden, C., Yu, H., Zheng, K. & Wang, S. 2015. Automated lesion detection on MRI scans using combined unsupervised and supervised methods. *BMC Medical Imaging*, vol. 15, no. 1, pp. 50.
- Guy, C. & Ffytche, D. 2005. *An Introduction to The Principles of Medical Imaging*, London, Imperial College Press.
- Hackmack, K., Paul, F., Weygandt, M., Allefeld, C. & Haynes, J. 2012. Multi-scale classification of disease using structural MRI and wavelet transform. *NeuroImage*, vol. 62, no. 1, pp. 48-58.
- Hagan, M., Demuth, H. & Beale, M. 1996. *Neural network design*, PWS Publishing Co.
- Haghighat, M., Zonouz, S. & Abdel-Mottaleb, M. 2015. CloudID: Trustworthy cloud-based and cross-enterprise biometric identification. *Expert Systems with Applications*, vol. 42, no. 21, pp. 7905-7916.
- Hall, M. A. 1999. Correlation-based feature selection for machine learning. PhD, University of Waikato, New Zealand.
- Hamel, L. 2009. *Knowledge Discovery with Support Vector Machines*, John Wiley & Sons, Inc.

- Han, J., Kamber, M. & Pei, J. 2011. *Data Mining: Concepts and Techniques*, Elsevier.
- Haralick, M., Shanmugam, K. & Dinstein, I. H. 1973. Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621.
- Hasan, A. & Meziane, F. 2016. Automated screening of MRI brain scanning using grey level statistics. *Computers & Electrical Engineering*, vol. 53, no., pp. 276–291.
- Hasan, A., Meziane, F. & Abd Kadhim, M. Automated Segmentation of Tumours in MRI Brain Scans. *Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC)*, 2016a, Rome, Italy. SCITEPRESS, pp. 55-62.
- Hasan, A., Meziane, F., Aspin, R. & Jalab, H. 2016b. Segmentation of Brain Tumors in MRI Images Using Three-Dimensional Active Contour without Edge. *Symmetry*, vol. 8, no. 11, pp. 132.
- Hasan, A. M., Meziane, F. & Jalab, H. A. Performance of grey level statistic features versus Gabor wavelet for screening MRI brain tumors: A comparative study. *International Conference on Information Communication and Management (ICICM)*, 2016c. IEEE, pp. 136-140.
- Haupt, R. Optimum population size and mutation rate for a simple real genetic algorithm that optimizes array factors. *Antennas and Propagation Society International Symposium, IEEE*, 2000. pp. 1034-1037.
- Haupt, R. & Haupt, S. 2004. *Practical Genetic Algorithms*, Canada, John Wiley & Sons.
- Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M. & Larochelle, H. 2017. Brain Tumor Segmentation with Deep Neural Networks. *Medical Image Analysis*, vol. 35, no. 1, pp. 18-31.
- Ho, S., Bullitt, L. & Gerig, G. Level-set evolution with region competition: automatic 3-D segmentation of brain tumors. *Proceedings of the 16th International Conference on Pattern Recognition, 2002 Canada. IEEE*, pp. 532-535.
- Howarth, P. & Ruger, S. 2005. Robust texture features for still-image retrieval. *IEEE Proceedings-Vision, Image and Signal Processing*, vol. 152, no. 6, pp. 868-874.
- Hu, Q. & Nowinski, W. L. 2003. A rapid algorithm for robust and automatic extraction of the midsagittal plane of the human cerebrum from neuroimages based on local symmetry and outlier removal. *NeuroImage*, vol. 20, no. 4, pp. 2153-2165.
- Hu, Y. & Hwang, J. 2001. *Handbook of Neural Network Signal Processing*, CRC press.
- Jähne, B. 2005. *Digital Image Processing*, Berlin : Springer.
- Jayalakshmi, T. & Santhakumaran, A. 2011. Statistical Normalization and Back Propagation for Classification. *International Journal of Computer Theory and Engineering*, vol. 3, no. 1, pp. 1793-8201.
- Jayasuriya, S. & Liew, A. Symmetry plane detection in neuroimages based on intensity profile analysis. *International Symposium on Information Technology in Medicine and Education (ITME)*, 2012, Australia, . IEEE, pp. 599-603.
- Jiang, J., Trundle, P. & Ren, J. 2010. Medical image analysis with artificial neural networks. *Computerized Medical Imaging and Graphics*, vol. 34, no. 8, pp. 617-631.
- Jin, L., Min, L., Wang, J., Fangxiang, W., Liu, T. & Yi, P. 2014. A survey of MRI-based brain tumor segmentation methods. *Tsinghua Science and Technology*, vol. 19, no. 6, pp. 578-595.

- Johnson, K. & Synovec, R. 2002. Pattern recognition of jet fuels: comprehensive GC×GC with ANOVA-based feature selection and principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, vol. 60, no. 1–2, pp. 225-237.
- Junck, L., Moen, J. G., Hutchins, G. D., Brown, M. B. & Kuhl, D. E. 1990. Correlation methods for the centering, rotation, and alignment of functional brain images. *Journal of nuclear medicine: official publication, Society of Nuclear Medicine*, vol. 31, no. 7, pp. 1220-1226.
- Kaal, E. C. & Vecht, C. J. 2004. The management of brain edema in brain tumors. *Current opinion in oncology*, vol. 16, no. 6, pp. 593-600.
- Kalbkhani, H., Shayesteh, M. & Zali-Vargahan, B. 2013. Robust algorithm for brain magnetic resonance image (MRI) classification based on GARCH variances series. *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 909-919.
- Karkavelas, G. & Tascos, N. 2011. Epidemiology, Histologic Classification and Clinical Course of Brain Tumors. DREVELEGAS, A. (ed.) *Imaging of Brain Tumors with Histological Correlations*. Springer Berlin Heidelberg.
- Kass, M., Witkin, A. & Terzopoulos, D. 1988. Snake: Active Contour Models. *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321-331.
- Kassner, A. & Thornhill, R. 2010. Texture analysis: a review of neurologic MR imaging applications. *American Journal of Neuroradiology*, vol. 31, no. 5, pp. 809-816.
- Kaus, M. R., Warfield, S. K., Nabavi, A., Black, P. M., Jolesz, F. A. & Kikinis, R. 2001. Automated segmentation of mr images of brain tumors. *Radiology*, vol. 218, no. 2, pp. 586-591.
- Khandani, M., Bajcsy, R. & Fallah, Y. Automated Segmentation of Brain Tumors in MRI Using Force Data Clustering Algorithm. In *International Symposium on Visual Computing, 2009, USA*. Springer Berlin Heidelberg, pp. 317-326.
- Kharrat, A., Karim, G., Ben Messaoud, M., Benamrane, N. & Abid, M. 2010. A Hybrid Approach for Automatic Classification of Brain MRI Using Genetic Algorithm and Support Vector Machine. *Leonardo Journal of Sciences*, vol. 1, no. 17, pp. 71-82.
- Khotanlou, H., Colliot, O., Atif, J. & Bloch, I. 2009. 3D brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models. *Fuzzy Sets and Systems*, vol. 160, no. 10, pp. 1457-1473.
- Kim, H. J. & Kim, W. 2012. Method of tumor volume evaluation using magnetic resonance imaging for outcome prediction in cervical cancer treated with concurrent chemotherapy and radiotherapy. *Radiation oncology journal*, vol. 30, no. 2, pp. 70-77.
- Klotz, A. 2013. 2D and 3D multiphase active contours without edges based algorithms for simultaneous segmentation of retinal layers from OCT images. MS.c, University of Texas at Austin, USA.
- Kong, W. K., Zhang, D. & Li, W. 2003. Palmprint feature extraction using 2-D Gabor filters. *Pattern recognition*, vol. 36, no. 10, pp. 2339-2347.
- Kuijf, H. J., Van Veluw, S. J., Geerlings, M. I., Viergever, M. A., Biessels, G. J. & Vincken, K. L. 2014. Automatic Extraction of the Midsagittal Surface from Brain MR Images using the Kullback–Leibler Measure. *Neuroinformatics*, vol. 12, no. 3, pp. 395-403.
- Lahmiri, S. & Boukadoum, M. 2013. Hybrid Discrete Wavelet Transform and Gabor Filter Banks Processing for Features Extraction from Biomedical Images. *Journal of Medical Engineering*, vol. 2013, no. 2013, pp.
- Larose, D. 2005. *Discovering Knowledge in Data An Introduction to Data Mining*, USA, John Wiley & Sons, Inc.

- Lauwers, L., Barbe, K., Van Moer, W. & Pintelon, R. Estimating the parameters of a Rice distribution: A Bayesian approach. International Instrumentation and Measurement Technology Conference, 2009, Singapore. IEEE, pp. 114-117.
- Lekutai, G. 1997. Adaptive Self-Tuning Neuro Wavelet Network Controllers. PhD, Virginia Polytechnic Institute and State University, USA.
- Li, S. & Ogihara, M. 2006. Using discriminant analysis for multi-class classification: an experimental investigation. Knowledge and Information Systems, vol. 10, no. 4, pp. 453-472.
- Liu, S. 2009. Symmetry and asymmetry analysis and its implications to computer-aided diagnosis: A review of the literature. Journal of Biomedical Informatics, vol. 42, no. 6, pp. 1056-1064.
- Liu, Y. & Collins, R. 1996. Automatic Extraction of the Central Symmetry (MidSagittal) Plane from Neuroradiology Images. Robotics Institute, Carnegie Mellon University, USA. Available: http://vision.cse.psu.edu/people/chenpingY/paper/liu_yanxi_1996_1.pdf
- Liu, Y., Collins, R. & Rothfus, W. Automatic bilateral symmetry (midsagittal) plane extraction from pathological 3D neuroradiological images. Proceedings of Medical Imaging 1998: Image Processing conference, 1998 USA pp. 1528-1539.
- Liu, Y., Muftan, M., Das, T., Bai, L., Robson, K. & Auer, D. 2012. Classification of MR Tumor Images Based on Gabor Wavelet Analysis. Journal of Medical and Biological Engineering, vol. 32, no. 1, pp. 22-28.
- Loizou, C. C., Pantziaris, M., Seimenis, I. & Pattichis, C. S. Brain MR image normalization in texture analysis of multiple sclerosis. 9th International Conference on Information Technology and Applications in Biomedicine, 2009, Larnaca. IEEE, pp. 1-5.
- M. Hall. 1999. Correlation-based Feature Selection for Machine Learning. PhD, University of Waikato, New Zealand.
- Manly, B. 1994. Multivariate Statistical Methods A primer, CRCpress
- Materka, A. & Strzelecki, M. 1998. Texture Analysis Methods – A Review. Technical University of Lodz, Institute of Electronics, COST B11 report , Brussels. Available: http://www.eletel.p.lodz.pl/programy/cost/pdf_1.pdf
- Matlab 2013. The Math Works kit.
- Mayfield Clinic for Brain & Spine Institute. 2013. anatomy of the brain [Online]. Available: www.mayfieldclinic.com.
- Mcinerney, T. & Terzopoulos, D. 2009. Chapter 8 - Deformable Models A2 - BANKMAN, ISAAC N. Handbook of Medical Image Processing and Analysis (Second Edition). Burlington: Academic Press.
- Mcrobbie, D. W., Moore, E. A., Graves, M. J. & Prince, M. R. 2007. MRI from Picture to Proton, Cambridge university press.
- Mechtler, L. 2009. Neuroimaging in Neuro-Oncology. Neurologic Clinics, vol. 27, no. 1, pp. 171-201.
- Menze, B. H., Van Leemput, K., Lashkari, D., Weber, M.-A., Ayache, N. & Golland, P. A generative model for brain tumor segmentation in multi-modal images. In International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2010, China. Springer Berlin Heidelberg, pp. 151-159.
- Menze, H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M. A., Arbel, T., Avants, B. B., Ayache, N., Buendia, P., Collins, D. L., Cordier, N., Corso, J. J., Criminisi, A., Das,

- T., Delingette, H., Demiralp, C., Durst, C. R., Dojat, M., Doyle, S., Festa, J., Forbes, F., Geremia, E., Glocker, B., Golland, P., Xiaotao, G., Hamamci, A., Iftekharuddin, K. M., Jena, R., John, N. M., Konukoglu, E., Lashkari, D., Mariz, J. A., Meier, R., Pereira, S., Precup, D., Price, S. J., Riklin Raviv, T., Reza, S. M. S., Ryan, M., Sarikaya, D., Schwartz, L., Hoo-Chang, S., Shotton, J., Silva, C. A., Sousa, N., Subbanna, N. K., Szekely, G., Taylor, T. J., Thomas, O. M., Tustison, N. J., Unal, G., Vasseur, F., Wintermark, M., Dong Hye, Y., Liang, Z., Binsheng, Z., Zikic, D., Prastawa, M., Reyes, M. & Van Leemput, K. 2015. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993-2024.
- Mihelič, F. & Žibert, J. 2008. *Speech Recognition: Technologies and Applications*, Vienna, Austria, I-Tech.
- Mikulka, J. & Gescheidtov, E. An Improved Segmentation of Brain Tumor, Edema and Necrosis. *Progress In Electromagnetics Research Symposium Proceedings, Taipei, 2013*. pp. 25-28.
- Moghaddasi, Z., Jalab, H., Md Noor, R. & Aghabozorgi, S. 2014. Improving RLRN Image Splicing Detection with the Use of PCA and Kernel PCA. *The Scientific World Journal*, vol. 2014, no. 2014, pp. 10.
- Mohammadi, S. M., Helfroush, M. S. & Kazemi, K. 2012. Novel shape-texture feature extraction for medical X-ray image classification. *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 1(B), pp. 659-676.
- Mohan, J., Krishnaveni, V. & Guo, Y. 2014. A survey on the magnetic resonance image denoising methods. *Biomedical Signal Processing and Control*, vol. 9, no., pp. 56-69.
- Morris, T. *Image Processing with MATLAB*. Manchester University, UK. Available: <https://studentnet.cs.manchester.ac.uk/ugt/COMP27112/doc/matlab.pdf>
- Mortazavi, D., Kouzani, A. Z. & Soltanian-Zadeh, H. 2012. Segmentation of multiple sclerosis lesions in MR images: a review. *Neuroradiology*, vol. 54, no. 4, pp. 299-320.
- Nabizadeh, N. 2015. *Automated Brain Lesion Detection and Segmentation Using Magnetic Resonance Images*. PhD, University of Miami, USA.
- Nabizadeh, N. & Kubat, M. 2015. Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features. *Computers & Electrical Engineering*, vol. 45, no., pp. 286-301.
- Nagarajan, M., Huber, M., Schlossbauer, T., Leinsinger, G., Krol, A. & Wismüller, A. 2013. Classification of Small Lesions in Breast MRI: Evaluating The Role of Dynamically Extracted Texture Features Through Feature Selection. *Journal of Medical and Biological Engineering*, vol. 33, no. 1, pp.
- Nailon, W. H. 2010. *Texture analysis methods for medical image characterisation*, INTECH Open Access Publisher.
- Naji, S., Zainuddin, R. & Jalab, H. 2012. Skin segmentation based on multi pixel color clustering models. *Digital Signal Processing*, vol. 22, no. 6, pp. 933-940.
- Naji, S., Zainuddin, R., Kareem, S. & Jalab, H. 2013. Detecting Faces in Colored Images Using Multi-skin Color Models and Neural Network with Texture Analysis. *Malaysian Journal of Computer Science*, vol. 26, no. 2, pp.
- Negnevitsky, M. 2005. *Artificial intelligence: a guide to intelligent systems*, Pearson Education.
- Nixon, M. & Aguado, A. 2008. *Feature Extraction Image Processing, Second Edition*, Academic Press.

- Nowak, R. 1999. Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Transactions on Image Processing*, , vol. 8, no. 10, pp. 1408-1419.
- Ortiz, A., Górriz, M., Ramírez, J., Salas-González, D. & Llamas-Elvira, M. 2013. Two fully-unsupervised methods for MR brain image segmentation using SOM-based strategies. *Applied Soft Computing*, vol. 13, no. 5, pp. 2668-2682.
- Özkan, C. & Erbek, F. 2003. The comparison of Activation Functions for Multispectral Landsat TM Image Classification. *Photogrammetric Engineering & Remote Sensing*, vol. 69, no. 11, pp. 1225-1234.
- Padma, A. & Sukanesh, R. 2011. Automatic Classification and Segmentation of Brain Tumor in CT Images using Optimal Dominant Gray level Run length Texture Features. *International Journal of Advanced Computer Sciences and Applications*, vol. 2, no. 10, pp. 53-59.
- Panchal, G., Ganatra, A., Kosta, Y. & Pancha, D. 2011. Behaviour Analysis of Multilayer Perceptrons with Multiple Hidden Neurons and Hidden Layers. *International Journal of Computer Theory and Engineerin*, vol. 3, no. 2, pp. 332.
- Pantelis, G. 2010. Computer Assisted Diagnosis of Brain Tumors based on Statistical Methods and Pattern Recognition Techniques. PhD, University of Patras, Greece.
- Parisot, S. 2013. Understanding, Modeling and Detecting Brain Tumors: Graphical Models and Concurrent Segmentation/Registration methods. PhD, Ecole Centrale Paris, France.
- Petrou, M. 2011. Texture in Biomedical Images. DESERNO, T. (ed.) *Biomedical Image Processing*. Springer Berlin Heidelberg.
- Petrov, N. & Jordanov, I. Unsupervised Texture Image Classification using Self-Organizing Maps. *International conference on artificial intelligence (ICAI'11)*, 2011, USA. pp.
- Pham, D., Xu, C. & Prince, J. 1998. A survey of current methods in medical image segmentation. USA. Available: <http://www.academia.edu/download/26282147/pham.pdf>
- Pock, T., Cremers, D., Bischof, H. & Chambolle, A. An algorithm for minimizing the Mumford-Shah functional. *12th International Conference on Computer Vision 2009*, Switzerland. IEEE, pp. 1133-1140.
- Prastawa, M., Bullitt, E. & Gerig, G. 2009. Simulation of brain tumors in MR images for evaluation of segmentation efficacy. *Medical image analysis*, vol. 13, no. 2, pp. 297-311.
- Prastawa, M., Bullitt, E., Ho, S. & Gerig, G. 2004. A brain tumor segmentation framework based on outlier detection. *Medical image analysis*, vol. 8, no. 3, pp. 275-283.
- Prima, S. & Ourselin, S. 2002. Computation of the mid-sagittal plane in 3-D brain images. *IEEE Transactions on Medical Imaging*, vol. 21, no. 2, pp. 122-138.
- Pujar, J., Gurjal, P., Shambhavi, S. & Kunnur, K. 2010. Medical Image Segmentation based on Vigorous Smoothing and Edge Detection Ideology. *World Academy of Science, Engineering and Technology*, vol. 68, no., pp. 444-450.
- Qinggang, W., Yong, G., Bin, L., Qiuwen, Z. & Huawen, C. 2015. An active contour model based on fused texture features for image segmentation. *Neurocomputing*, vol. 151, no., pp. 1133-1141.
- Quinn, G. & Keough, M. 2002. *Experimental Design and Data Analysis for Biologists*.
- Rahim, M. S. M., Saba, T., Nayer, F. & Syed, A. Z. 2014. 3D texture features mining for MRI brain tumor identification. *3D Research*, vol. 5, no. 1, pp. 1-8.
- Ray, N., Greiner, R. & Murtha, A. 2008a. Using Symmetry to Detect Abnormalities in Brain MRI. *Computer Society of India Communications*, vol. 31, no. 19, pp. 7-10.

- Ray, N., Saha, N. & Hong, Z. Change Detection and Object Segmentation: A Histogram of Features-Based Energy Minimization Approach. Sixth Indian Conference on Computer Vision, Graphics & Image Processing, 2008b India. IEEE, pp. 628-635.
- Rogowska, J. 2009. Chapter 5 - Overview and Fundamentals of Medical Image Segmentation A2 - BANKMAN, ISAAC N. Handbook of Medical Image Processing and Analysis (Second Edition). Burlington: Academic Press.
- Rojas, R. 1996. Neural Networks a Systematic Introduction, Springer-Verlag, Berlin.
- Ross, T. J. 2009. Fuzzy logic with engineering applications, John Wiley & Sons.
- Rousseau, O. 2009. Geometrical Modeling of the Heart. Ph.D, University of Ottawa, Canada.
- Ruppert, G., Teverovskiy, L., Chen-Ping, Y., Falcao, X. & Yanxi, L. A new symmetry-based method for mid-sagittal plane extraction in neuroimages. International Symposium on Biomedical Imaging: From Nano to Macro, 2011 USA. IEEE, pp. 285-288.
- Russ, J. C. 1990. Image processing, Springer.
- Sachdeva, J., Kumar, V., Gupta, I., Khandelwal, N. & Ahuja, C. 2013. Segmentation, Feature Extraction and Multiclass Brain Tumor Classification. Swetswise, vol. 26, no. 6, pp. 1141-1150.
- Sachdeva, J., Kumar, V., Gupta, I., Khandelwal, N. & Ahuja, C. 2016. A package-SFERCB-“Segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors”. Applied Soft Computing, vol. 47, no., pp. 151-167.
- Sachdeva, J., Kumar, V., Gupta, I., Khandelwal, N. & Ahuja, C. K. 2012. A novel content-based active contour model for brain tumor segmentation. Magnetic Resonance Imaging, vol. 30, no. 5, pp. 694-715.
- Saha, B., Ray, N., Greiner, R., Murtha, A. & Zhang, H. 2012. Quick detection of brain tumors and edemas: A bounding box method using symmetry. Computerized Medical Imaging and Graphics, vol. 36, no. 2, pp. 95-107.
- Sanjuán, A., Price, C. J., Mancini, L., Josse, G., Grogan, A., Yamamoto, A. K., Geva, S., Leff, A. P., Yousry, T. A. & Seghier, M. L. 2013. Automated identification of brain tumors from single MR images based on segmentation with refined patient-specific priors. Frontiers in neuroscience, vol. 7, no., pp.
- Saritha, M., Paul Joseph, K. & Mathew, A. 2013. Classification of MRI brain images using combined wavelet entropy based spider web plots and probabilistic neural network. Pattern Recognition Letters, vol. 34, no. 16, pp. 2151-2156.
- Sato, M., Lakare, S., Wan, M., Kaufman, A. & Nakajima, M. A gradient magnitude based region growing algorithm for accurate segmentation. International Conference on Image Processing, 2000 Canada. IEEE, pp. 448-451.
- Schild, H. 1990. MRI made easy well almost, Germany.
- Schmid, P. 1999a. Segmentation and symmetry measure for image analysis: application to digital dermatoscopy. PhD, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, Switzerland.
- Schmid, P. 1999b. Segmentation of digitized dermatoscopic images by two-dimensional color clustering. Medical Imaging, IEEE Transactions on, vol. 18, no. 2, pp. 164-171.
- Smith, L. I. 2002. A tutorial on Principal Components Analysis. Cornell University, USA, vol. 51, no., pp. 52.
- Soille, P. 2003. Morphological Image Analysis: Principles and Applications, Springer-Verlag New York, Inc.

- Solomon, C. & Breckon, T. 2011. *Fundamentals of Digital Image Processing: A practical approach with examples in Matlab*, John Wiley & Sons.
- Song, H., Cui, X. & Sun, F. 2013. Breast tissue 3D segmentation and visualization on MRI. *Journal of Biomedical Imaging*, vol. 2013, no., pp. 20.
- Sonka, M., Hlavac, V. & Boyle, R. 2014. *Image Processing, Analysis and Machine Vision*, Cengage Learning.
- Stańczyk, U. & Jain, L. 2015. *Feature Selection for Data and Pattern Recognition*, Berlin Heidelberg, Springer.
- Tai, X.-C., Lie, K.-A., Chan, T. F. & Osher, S. *Image Processing Based on Partial Differential Equations. Proceedings of the International Conference on PDE-Based Image Processing and Related Inverse Problems, 2005 Oslo*. Springer Science & Business Media, pp.
- Takayanagi, Y., Takahashi, T., Orikabe, L., Mozue, Y., Kawasaki, Y., Nakamura, K., Sato, Y., Itokawa, M., Yamasue, H. & Kasai, K. 2011. Classification of first-episode schizophrenia patients and healthy subjects by automated MRI measures of regional brain volume and cortical thickness. *PloS one*, vol. 6, no. 6, pp. e21047.
- Talebi, M., Ayatollahi, A. & Kermani, A. 2010. Medical ultrasound image segmentation using genetic active contour. *Biomedical Science and Engineering*, vol. 4, no. 2, pp. 105-109.
- Tang, J., Alelyani, S. & Liu, H. 2014. Feature selection for classification: A review. *Data Classification: Algorithms and Applications*, vol., no., pp. 37.
- Tang, X. 1998. Texture information in run-length matrices. *IEEE transactions on image processing*, vol. 7, no. 11, pp. 1602-1609.
- Tantisatirapong, S. 2015. *Texture analysis of multimodal magnetic resonance images in support of diagnostic classification of childhood brain tumours*. PhD, University of Birmingham, UK.
- Thapaliya, K., Pyun, J., Park, C. & Kwon, G. 2013. Level set method with automatic selective local statistics for brain tumor segmentation in MR images. *Computerized Medical Imaging and Graphics*, vol. 37, no. 7, pp. 522-537.
- Tian, D. 2013. *A Review on Image Feature Extraction and Representation Techniques*. *International Journal of Multimedia and Ubiquitous Engineering*, vol. 8, no. 4, pp. 385-396.
- Toennies, K. D. 2012. *Guide to medical image analysis: methods and algorithms*, Springer Science & Business Media.
- Tohka, J. 2014. Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: A review. *World journal of radiology*, vol. 6, no. 11, pp. 855–864.
- Tonarelli, L. 2013. *Magnetic Resonance Imaging of Brain Tumor*. *Enterprises for Continuing Education.*, vol. 300, no., pp. 48116-0300.
- Tu, C., Chuang, L., Chang, J. & Yang, C. 2007. Feature Selection using PSO-SVM. *IAENG International Journal of Computer Science*, vol. 33, no. 1, pp.
- Tuzikov, A. V., Colliot, O. & Bloch, I. 2003. Evaluation of the symmetry plane in 3D MR brain images. *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2219-2233.
- Uc Davis Medical Center. 2016. *Troubleshooting Guide; Prep, Scan Errors and Artifacts*. University of California, USA. Available: <http://ucdir.ucdavis.edu/start/doc/troubleshooting.pdf>

- Udomchaiporn, A., Coenen, F., García-Fiñana, M. & Sluming, V. 2013. 3-D MRI Brain Scan Feature Classification Using an Oct-Tree Representation. *Advanced Data Mining and Applications*. Springer.
- Van Der Maaten, L. 2007. An introduction to dimensionality reduction using matlab. Report MICC 07-07, Universiteit Maastricht, Netherlands. Available: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.107.1327&rep=rep1&type=pdf>
- Verma, R., Zacharaki, E. I., Ou, Y., Cai, H., Chawla, S., Lee, S.-K., Melhem, E. R., Wolf, R. & Davatzikos, C. 2008. Multiparametric tissue characterization of brain neoplasms and their recurrence using pattern classification of MR images. *Academic radiology*, vol. 15, no. 8, pp. 966-977.
- Wallisch, P., Lusignan, M., Benayoun, M., Baker, T., Dickey, A. & Hatsopoulos, N. 2014. *MATLAB for neuroscientists: an introduction to scientific computing in MATLAB*, Academic Press.
- Wang, Q. 2012. Kernel principal component analysis and its applications in face recognition and active shape models. *arXiv preprint arXiv:1207.3538*, vol., no., pp.
- Watt, A. & Policarpo, F. 1998. *The computer image*, Addison Wesley.
- William, K. 2001. *Digital Image Processing*, Canada, Wiley-Interscience.
- Wilson, J. & Ritter, G. 2000. *Handbook of Computer Vision Algorithms in Image Algebra*, CRC Press, Inc.
- Witten, I., Frank, E. & Hall, M. 2011. *Data Mining Practical Machine Learning Tools and Techniques*, Elsevier.
- Xi-Ping, L., Jie, T. & Yao, L. 2002. An Algorithm for Segmentation of Medical Image Series Based on Active Contour Model. *Journal of Software*, vol. 13, no. 6, pp. 1050-1058.
- Xu, C., Pham, D. L. & Prince, J. L. 2000. Image segmentation using deformable models. *Handbook of medical imaging*, vol. 2, no., pp. 129-174.
- Xu, C. & Prince, J. L. 1998. Snakes, shapes and gradient vector flow. *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359-369.
- Xu, S. & Chen, L. A Novel Approach for Determining the Optimal Number of Hidden Layer Neurons for FNN's and Its Application in Data Mining. *5th International Conference on Information Technology and Applications (ICITA)*, 2008, Australia. pp. 683-686.
- Yang, X., Tridandapani, S., Beitler, J., Yu, D., Yoshida, E., Curran, W. & Liu, T. 2012. 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*, vol. 39, no. 9, pp. 5732-5739.
- Zacharaki, E. I., Wang, S., Chawla, S., Soo Yoo, D., Wolf, R., Melhem, E. R. & Davatzikos, C. 2009. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magnetic Resonance in Medicine*, vol. 62, no. 6, pp. 1609-1618.
- Zhang, G., Eddy Patuwo, B. & Hu, M. 1998. Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting*, vol. 14, no. 1, pp. 35-62.
- Zhang, H., Fritts, J. E. & Goldman, S. A. 2008. Image segmentation evaluation: A survey of unsupervised methods. *computer vision and image understanding*, vol. 110, no. 2, pp. 260-280.
- Zhang, Y., Brady, J. M. & Smith, S. M. 2001. An HMRF-EM Algorithm for Partial Volume Segmentation of Brain MRI FMRIB Technical Report TR01YZ1. Oxford Centre for

Functional Magnetic Resonance Imaging of the Brain. Available:
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.65.2102&rep=rep1&type=pdf>
Zulpe, N. & Pawar, V. 2012. GLCM Textural Features for Brain Tumor Classification.
International Journal of Computer Science Issues (IJCSI), vol. 9, no. 3, pp. 354-359.

APPENDIX A

To measure noise sensitivity of BBBGA, Gaussian noise of zero mean with different $\sigma = 0, 0.1, 0.2, 0.3$ and 0.4 are used to contaminate the pathological T2-w MRI slices (Fig. A.1, Fig A.2, Fig. A.3, Fig. A.4 and Fig. A.5). It is noted that the efficacy of BBBGA to identify the location of brain tumour decreases with increasing amount of noise in MRI scan. By increasing the amount of noise the overall brightness of MRI slices increases and the spatial scale of normal tissue becomes relatively similar to normal variability in the pathology. As shown in Fig. 7.15, the FP in this example represents the part of the generated 3D-box by BBBGA, which is located outside the brain tumour boundary. Therefore, the FP at noise-free MRI scan is relatively small, meaning that the major area of the generated 3D-box is located inside the boundary of the brain tumour. Increasing the amount of noise in MRI scan leads to increase the normal anatomy tissue that is covered by 3D-box, resulting in increasing the probability of the error in choosing the seed point to initialize the segmentation process.

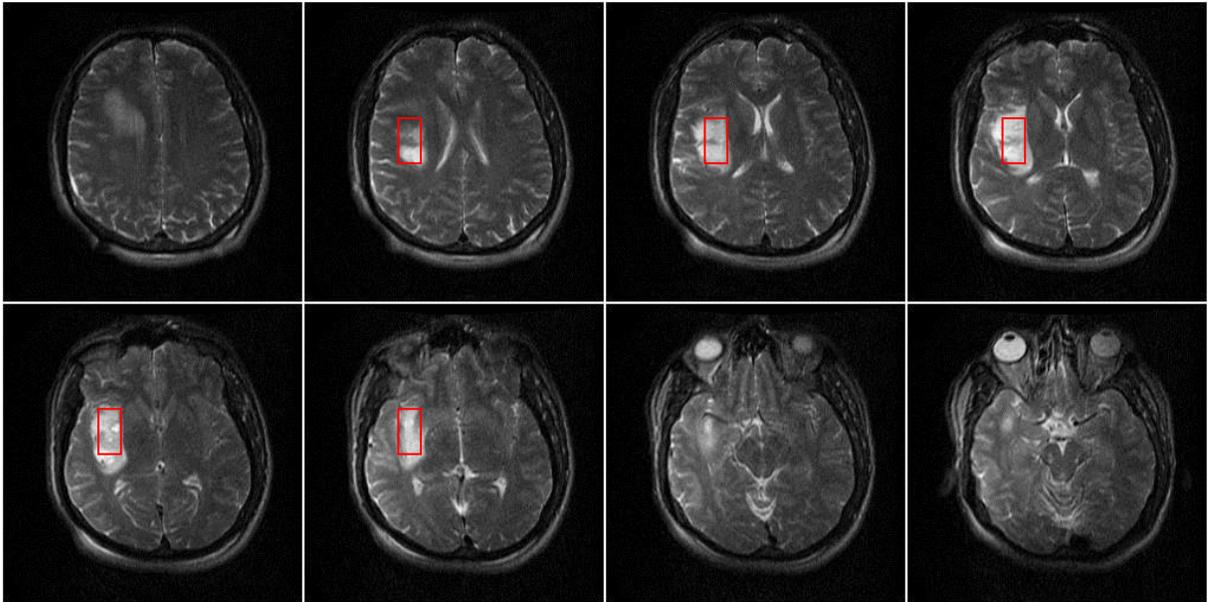


Figure A.1: Result of the BBBGA implementation on MRI brain scan of a pathological patient when there is no additive Gaussian noise ($\sigma = 0$).

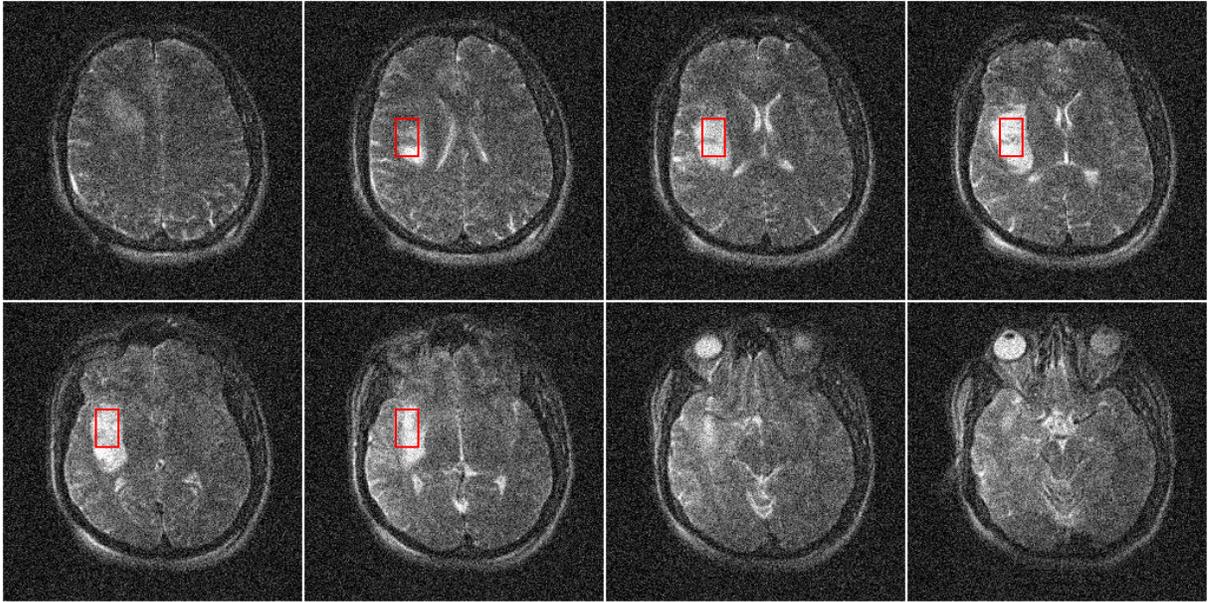


Figure A.2: Result of the BBBGA implementation on MRI brain scan of a pathological patient when there is additive Gaussian noise ($\sigma = 0.1$).

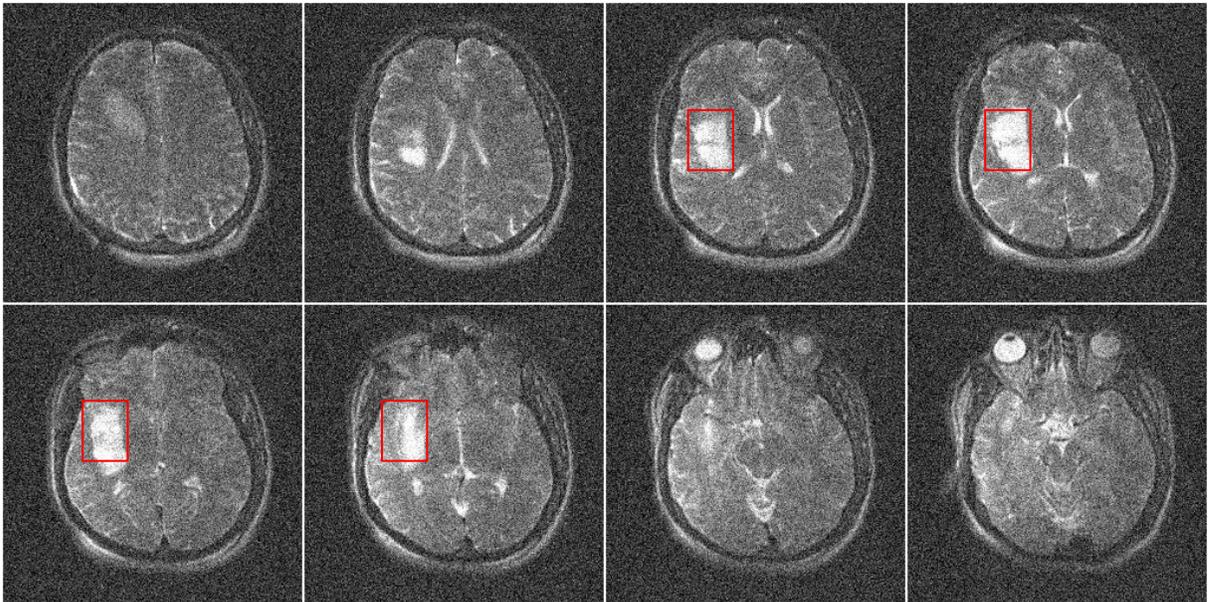


Figure A.3: Result of the BBBGA implementation on MRI brain scan of a pathological patient when there is additive Gaussian noise ($\sigma = 0.2$).

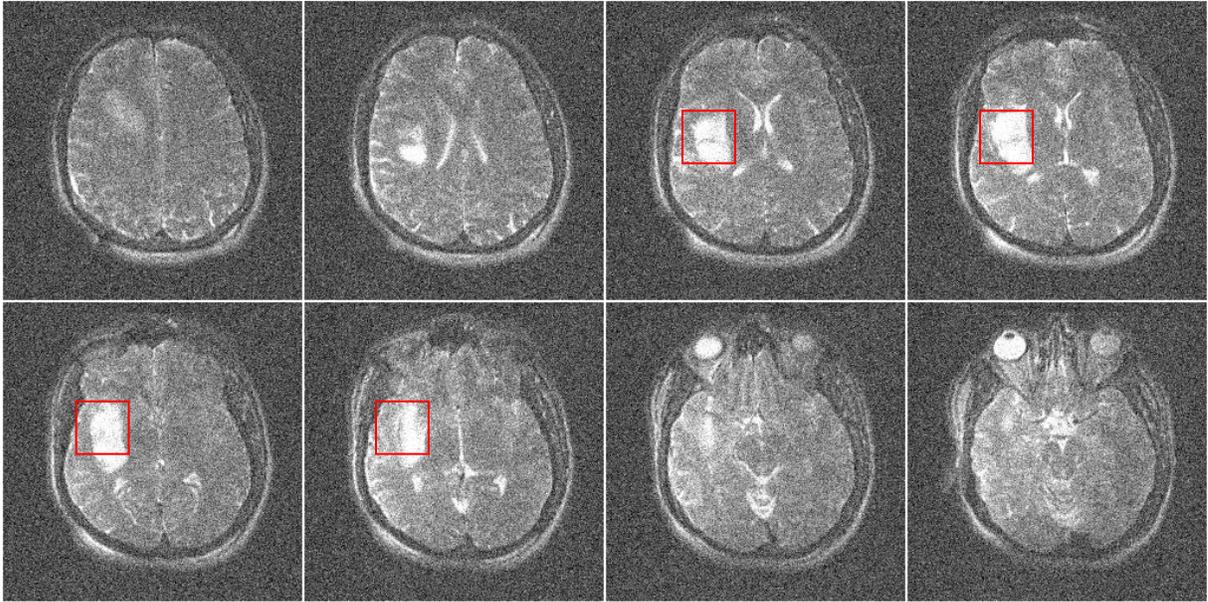


Figure A.4: Result of the BBBGA implementation on MRI brain scan of a pathological patient when there is additive Gaussian noise ($\sigma = 0.3$).

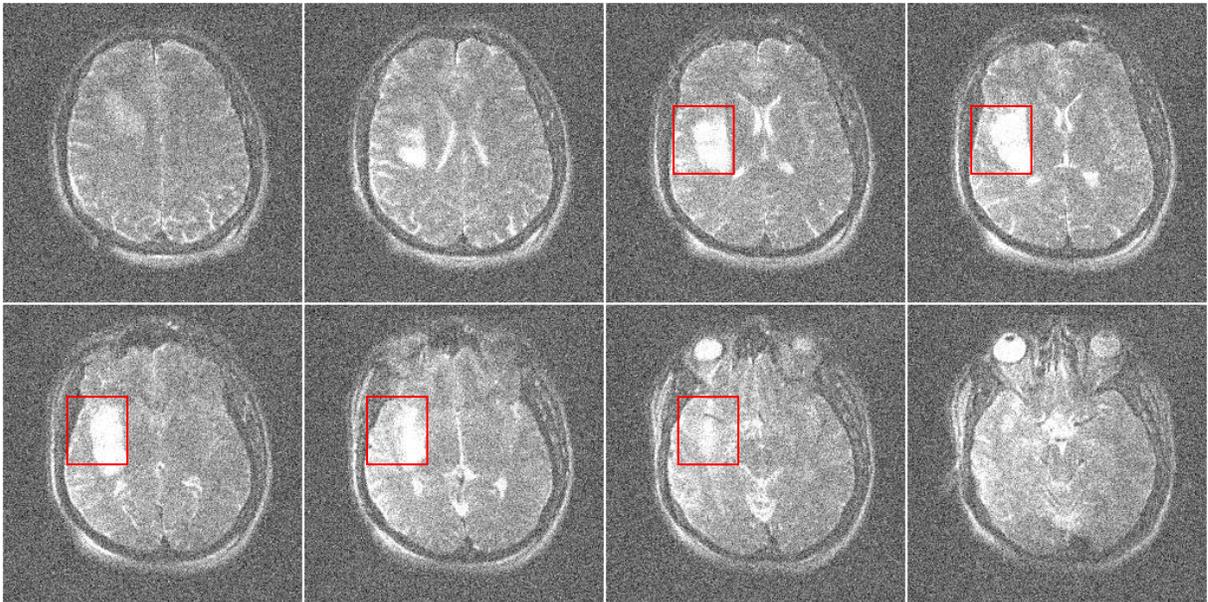


Figure A.5: Result of the BBBGA implementation on MRI brain scan of a pathological patient when there is additive Gaussian noise ($\sigma = 0.4$).