Computer-Aided Detection of Epileptogenic Lesions Based on Brain MRI Image Analysis

by

Huiquan Wang

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Signal and Image Processing

Department of Electrical and Computer Engineering

University of Alberta

Abstract

The invention of magnetic resonance imaging (MRI) has advanced the diagnosis of diseases dramatically as it is capable of producing high-quality images of soft tissues non-invasively and safely. However, the high volume of MRI images also means heavy workloads for physicians, which results in slow diagnosis and even misdiagnosis. As a consequence, computer-aided detection (CAD) systems, which can process the input images by computer and generate diagnosis results for radiologists to improve the diagnostic efficiency and accuracy, are in high demand. Some CAD systems for brain disorders, such as brain tumors, Alzheimer's disease, and multiple sclerosis, have already been proposed and actively researched. But there is no CAD system for screening epileptogenic lesions even though there is a large population of epilepsy patients worldwide. In this thesis, several automated detection techniques that can aid build a CAD system for epileptogenic lesions are developed.

The contributions of this thesis are threefold. First, an automated detection technique for cavernous malformation based on MRI image analysis is proposed. The technique is a coarse to fine framework. After obtaining the candidate cavernous malformation regions using skull stripping and template matching, image features including texture, the histogram of oriented gradients, and local binary patterns of each candidate are extracted and then classified using support vector machines to exclude false detections. Second, an automated detection technique for mesial temporal sclerosis (MTS) is proposed. After the segmentation and 3D reconstruction of the hippocampus,

the hippocampal volume, shape, and cerebrospinal fluid features are calculated. Two support vector machines are then used to detect and lateralize MTS. Third, an automated detection technique for focal cortical dysplasia is proposed using MRI images and deep learning. After bias field correction, intensity normalization, and registration with a brain atlas, cortical patches are extracted and fed to a deep convolutional neural network with five convolutional layers, a max pooling layer, and two fully-connected layers. Image patches with focal cortical dysplasia are detected after classification and post-processing.

The proposed techniques are evaluated thoroughly using both a publicly available MRI dataset and images obtained from the University of Alberta Hospital. Experimental results indicate that the proposed techniques could provide superior performance compared with existing methods in the literature, thus showing potential in assisting neuroradiologists in the detection of epileptogenic lesions using brain MRI analysis.

To my parents Youping Wang and Dexiang Wu,
my wife Yan Wang, and my daughter Emily Wang
for their endless support.

Acknowledgements

First of all, I would like to express my very great appreciation to my supervisor, Dr. Mrinal Mandal, for his patient guidance, enthusiastic encouragement, and continuous support throughout my Ph.D. studies at the University of Alberta. He is a knowledgeable, kind and considerate professor, and could always provide valuable advice and constructive suggestions for my research.

I would also like to express my gratitude to my co-supervisor, Dr. Nizam Ahmed, for teaching me medical knowledge about epilepsy, identifying MRI images, and giving me valuable suggestions in my research. The works in this thesis could not be conducted without his help.

I want to express my sincere thanks to my committee members for devoting their precious time in reviewing my candidacy report and dissertation, as well as providing their insightful comments for my research.

I appreciate the financial support from China Scholarship Council (CSC) in the first four years of my Ph.D. study. I am particularly grateful for the help from my labmates in Multimedia Computing and Communications Laboratory (MCCL): Hongming Xu, Yue Li, Bingjie Ma, Salah Al-Heejawi, Lina Liu and many others. Their companionship made my Ph.D. life more colorful.

Last but not least, I would like to thank my parents, my wife, and daughter for their endless love, encouragement, and support.

Table of Contents

1 Introduction **1**

1.1 Epilepsy 1

1.2 MRI-Based Diagnosis 3

1.3 Problem Statements and Motivation 7

1.4 Major Contributions of the Thesis 8

1.5 Organization of the Thesis 10

2 Literature Review **11**

2.1 Preprocessing 11

2.1.1 Smoothing 12

2.1.2 Inhomogeneity Correction 12

2.1.3 Intensity Standardization 13

2.1.4 Registration 13

2.2 Segmentation 14

2.2.1 Thresholding 14

2.2.2 Clustering 14

2.2.3 Active Contour Models 15

2.2.4 Atlas-Based Segmentation 16

2.3 Feature Extraction 16

2.3.1 First-order Textures 17

2.3.2 Histogram of Oriented Gradient 17

2.3.3 Local Binary Pattern 18

2.3.4 The Discrete Wavelet Transform 18
2.4 Classification 19
2.4.1 K-Nearest Neighbors 19
2.4.2 Logistic Regression 19
2.4.3 Support Vector Machines 20
2.4.4 Artificial Neural Networks 21
2.4.5 Deep Learning 22
2.4.6 Performance Evaluation Metrics 22
2.5 CAD Systems 24
2.6 Summary 27

3 Automated Detection of Cavernous Malformation 28
3.1 Introduction 29
3.2 Materials 32
3.3 Proposed Method 33
3.3.1 Skull Stripping 34
3.3.2 Template Matching 38
3.3.3 Feature Extraction 40
3.3.4 Classification 43
3.4 Experimental Results 46
3.4.1 Efficiency of Skull Stripping 46
3.4.2 System Performance 47
3.4.3 Computation Time 52
3.5 Conclusion 52

4 Automated Detection of Mesial Temporal Sclerosis 54
4.1 Introduction 55
4.2 Materials 58
4.3 Methods 58

4.3.1 Brain and Hippocampus Segmentation 59
4.3.2 Feature Extraction 60
4.3.3 Classification 63
4.3.4 Evaluation 64
4.3.5 Performance Comparison 65
4.3.6 Feature Significance 65
4.4 Experimental Results 68
4.4.1 Effectiveness of PCA 68
4.4.2 Classification Models 69
4.4.3 System Performance 71
4.4.4 Computation Time 74
4.4.5 Discussion 75
4.5 Conclusion 77

5 Automated Detection of Focal Cortical Dysplasia 78

5.1 Introduction 79
5.2 Materials 83
5.3 Methods 83
5.3.1 Preprocessing 84
5.3.2 Patch Extraction 84
5.3.3 Deep Learning Classification 87
5.3.4 Post-processing 92
5.4 Experimental Results 93
5.4.1 Implementation 93
5.4.2 Evaluation of the CNN 94
5.4.3 Evaluation of the System 97
5.5 Conclusion 101

6 Conclusions and Future Work **102**

6.1 Conclusions 102

6.2 Resulting Publications 104

6.3 Future Work 105

6.3.1 More Epileptogenic Lesions 105

6.3.2 Deep Learning in Epilepsy 106

6.3.3 Multiple Imaging Modalities 107

6.3.4 Accurate Segmentation of Epileptogenic Lesions 107

References **108**

List of Tables

2.1 First-order texture features. 17

2.2 Related CAD systems based on brain MRI image analysis. BT: brain tumor, AD: Alzheimer's disease, MS: multiple sclerosis, Dis.: disease, Cla.:classifier, Acc.: accuracy reported. 24

3.1 Performance comparison of skull stripping methods. 47

3.2 Parameters of training and testing. 49

3.3 Performance based on feature dimensionality 51

3.4 Performance comparison. 52

4.1 Statistical analysis of hippocampal volume and CSF-ratio features between groups (NC, LMTS, and RMTS). SD: standard deviation. 66

4.2 Performance comparison for different classifiers. Note: the values in parentheses are the standard errors estimated using binomial distributions. 70

4.3 Confusion matrix for LSVM and RSVM. Act.: Actual, Pred.: Prediction. 72

4.4 Performance comparison for different techniques. Sen.: sensitivity; Spe.: specificity; Acc.: accuracy. 74

5.1 Details of the proposed CNN architecture with 5 convolutional layers, 3×3 filters, and scale parameter $k = 4$. 88

5.2 Performance of different CNN architectures. Acc.: accuracy. . 95
5.3 Confusion matrix for FCD detection with 30 MR images. . . . 98
5.4 Performance comparison with existing techniques for FCD detection. Spe.: specificity, sen.: sensitivity, acc: accuracy 100
5.5 Performance comparison with existing techniques for FCD segmentation. 101

List of Figures

1.1 Magnetic resonance imaging. (a) MRI scanner cutaway [1], (b) a perturbed proton in magnetic field. 5

1.2 Brain MRI images. (a)(b)(c) are T1W images in sagittal, coronal, and axial view, respectively. (d)(e)(f) are T2W images in sagittal, coronal and axial view, respectively. Note that (c) is labeled with cerebrospinal fluid, white matter, gray matter, scalp, and skull. 5

1.3 Three examples of cavernous malformation in T2W axial plane images. (a)(b)(c) are upper, middle, and lower level brain slices with cavernomas (in square boxes). (d)(e)(f) are the enlarged cavernoma regions. H: hemosiderin, V: vessels. 6

1.4 Mesial temporal sclerosis. (a) T2W image, (b) the corresponding FLAIR image. 7

1.5 Focal cortical dysplasia in T1W images. 7

1.6 The framework of the proposed CAD system for epileptogenic lesions. 9

2.1 General framework of a CAD system. 12

2.2 Intensity inhomogeneity in brain MRI image. Note that the top parts of the brain area have lower intensities compared to the bottom parts as shown in the inhomogeneity field. 13

2.3 LBP value for the center pixel. 18

2.4 A three-layer ANN. 21

3.1 Three examples of cavernous malformations in T2W images. Top row: original images with cavernomas in yellow boxes. Bottom row: enlarged cavernoma regions. 30

3.2 Six samples of cavernous malformations. 33

3.3 Schematic of the proposed CAD system for cavernous malformations. 34

3.4 Skull stripping results of one slice of a 3D image. (a) An input T2W slice, (b) input slice overlaid with the brain mask (in green) after registration, (c) refined brain contour (shown by the yellow boundary), (d) extracted brain. 35

3.5 Templates used in the proposed technique. 39

3.6 Distribution of NCC values. (a) Maximum NCC values for each cavernoma, (b) the plot of the NCC cumulative distribution function. 40

3.7 The cell and block division in an image window. b and c refer to block and cell, respectively. 42

3.8 LBP value for the the center pixel. 43

3.9 Candidate region labeling. (a) The ground truth boundary (red) of a cavernoma, (b) a candidate region (yellow) is labeled as CM, (c) a candidate region (yellow) is labeled as non-CM. . 44

3.10 ROC for classification by SVMs. 46

3.11 Boxplot of dice coefficients for all patients. Red lines denote the median dice coefficient and the '+'s denote the outliers. . . . 48

3.12 An example of skull stripping. (a) The input image, (b) extracted brain (in yellow boundary) using atlas-based method, (c) extracted brain (in yellow boundary) using morphology-based method. Note that in (c), the cavernoma region (shown in red box) is not included in the brain area. 48
3.13 Two qualitative results of the proposed system. (a) Two input images corresponding to the upper level (top row) and lower level (bottom row) brain slices, (b) extracted brain after skull stripping, (c) candidate cavernoma regions in yellow boxes after template matching, (d) detected cavernomas (in red boxes) after SVM classification. 49
3.14 Performance of the CAD system for all patients. (a) Sensitivity, (b) specificity, and (c) accuracy 50
3.15 Cumulative variance plot. 51

4.1 Left and right hippocampi in 3D. 55
4.2 Two T1-weighted MRI coronal slices with hippocampal volume loss indicated by white arrows. (a) Left side MTS, (b) right side MTS. Note that the lost hippocampal volume is occupied by cerebrospinal fluid (CSF) and appears black in the indicated regions. 56
4.3 Example of an MR image with manually segmented hippocampus in three view planes. (a) Sagittal plane, (b) coronal plane, and (c) transverse plane. 59
4.4 The proposed MTS detection method. 59
4.5 Segmentation results of one brain MR image. (a) Input image, (b) CSF, (c) gray matter, and (d) white matter. 60

4.6 Hippocampus extraction results. (a) Extracted hippocampal area (the left and right hippocampus are in pink and red, respectively) in the coronal plane, (b) the mesh model for the right hippocampus. 60
4.7 Bar graph of mean volumetric and CSF-ratio features for three image groups (NC, LMTS, and RMTS). (a) Normalized hippocampal volume, (b) CSF-ratio, (c) hippocampal and CSF-ratio asymmetry index. Error bars are standard deviations. The scatter points (in blue color) show actual values of all subjects in each group. 67
4.8 Shape deformations. (a) Mean right hippocampal shape of NC images, (b) the right hippocampal shape deformation for the RMTS group, (c) mean left hippocampal shape of NC images, (d) the left hippocampal shape deformation for the LMTS group. Note that the color stands for the t-statistic values. . . 68
4.9 PCA for left MTS detection. (a) Performance plot using different number of principal components, (b) scree plot of the percent variability explained by each principal component. . . 69
4.10 Bar plot of SVM weights. (a) LSVM, (b) RSVM. 71
4.11 The two miss detections. (a) Right MTS indicated by the white arrow, (b) ground truth [64] of right hippocampus (in the white contour) for (a), (c) segmentation results of right hippocampus for (a), (d) left MTS pointed by the white arrow, (e) ground truth [64] of left hippocampus (in the white contour) for (d), (f) segmentation results of left hippocampus for (d). 73
4.12 ROC curves for LSVM and RSVM 73
4.13 MTS detection using different features. (a) Left MTS detection, (b) right MTS detection. 74

5.1 T1-weighted MR images. (a) An axial slice of a healthy brain, (b) a FCD slice with blurred gray/white matter boundary, (c) a FCD slice with increased cortical thickness (typical values of cortical thickness for adults are between 1.5 mm and 3.0 mm [133]). 80
5.2 Schematic of the proposed FCD detection method. 84
5.3 Patch masks extraction on one axial slice of GMPM. (a) The axial slice divided into horizontal strips of height h, (b) one horizontal strip, (c) line plot of the Radon transform of (b), (d) two external patches (the green boxes), and (e) two middle patches (the blue boxes) along the inter-hemisphere fissure. Note that all patch masks have the same size of $h \times w$. 86
5.4 Patches extracted on one axial slice. 86
5.5 Schematic of the proposed CNN architecture (with 5 convolutional layers). 87
5.6 Fully-connected layers. 91
5.7 A patch (P) and its four neighbors (N_1-N_4). 93
5.8 The ROC curve for the proposed CNN. 95
5.9 Loss and accuracy plots during the network training process. . 96
5.10 The first layer filters. 97
5.11 An example of FCD detection. (a) an axial slice with FCD lesion labeled in red, (b) patch extraction results, (c) classification results and numbers stand for probabilities of being FCD patches, (d) detection results mapped onto the inflated cortical surface (shown in yellow). 99
5.12 The missed FCD detection. (a) an axial slice with the enlarged FCD lesion, (b) patch probabilities (in magenta) of being FCD. 100

List of Abbreviations

AD	Alzheimer's disease
AVM	arteriovenous malformation
CAD	computer-aided detection
CM	cavernous malformation
CNN	convolutional neural network
CSF	cerebrospinal fluid
CT	computed tomography
DWT	discrete wavelet transform
EEG	electroencephalogram
FCD	focal cortical dysplasia
FLAIR	fluid-attenuated inversion recovery
gLoG	generalized Laplacian of Gaussian
GMC	gray matter concentration
GME	gray matter extension
GPU	graphics processing unit
GWJ	gray matter white matter junction
HOG	histogram of oriented gradients
KNN	K-Nearest Neighbors
LBP	local binary pattern
LR	logistic regression
MRI	magnetic resonance imaging
MS	multiple sclerosis

MTS	mesial temporal sclerosis
NCC	normalized cross correlation
PCA	principal component analysis
PET	positron-emission tomography
PVE	partial volume estimation
ReLU	rectified linear units
ROC	receiver operating characteristic
ROI	region of interest
SBC	surface-based classification
SGD	stochastic gradient descent
SPM	statistical parametric mapping
T1W	T1-weighted
T2W	T2-weighted
TLE	temporal lobe epilepsy
TM	template matching
VBM	voxel-based morphometry
VNS	vagus nerve stimulation
VOI	volume of interest

List of Symbols

$\bar{\boldsymbol{S}}$	the mean hippocampal shape
σ	the scale factor in batch normalization
β	the offset factor in batch normalization
γ	the output probability of softmax function
ξ_i	the ith slack variable in SVM
∇f	the gradient of the function f
$\boldsymbol{\phi}$	texture feature vector
θ	orientation of the gLoG kernel
τ	predefined threshold
σ_x, σ_y	scales of the gLoG kernel
H_b	the block gradient orientation histogram
H_c	the cell gradient orientation histogram
B_0	the magnetic field
C	the regularization parameter in SVM
M	the proton magnetic moment
M_{xy}	the transverse moment
M_z	the longitudinal moment
N_{FN}	the number of false negatives
N_{FP}	the number of false positives
N_{TN}	the number of true negatives
N_{TP}	the number of true positives
z_i	a pixel intensity

Chapter 1

Introduction

Medical imaging techniques (X-ray, ultrasound, MRI, *etc.*) are playing essential roles in the diagnosis of patients at the hospital. Radiologists and clinicians usually have to inspect and analyse a large number of images obtained using these imaging techniques in a short time. To lessen the burden of doctors and assist them in the interpretation of medical images, some automaged detection methods for various diseases (breast cancer, lung cancer, coronary artery disease, *etc.*) have already been proposed. The objective of this thesis is to develop a comprehensive computer-aided detection (CAD) system for epileptogenic lesions including cavernous malformation, mesial temporal sclerosis, and focal cortical dysplasia, based on brain MRI image analysis. In this chapter, a brief background on epilepsy along with its diagnosis based on MRI are first introduced. The research motivations and problems to solve are then presented. After that, the major contributions of this thesis are presented followed by the organization of a review of this thesis.

1.1 Epilepsy

Epilepsy is a brain disease characterized by unprovoked seizures or epileptic features such as abnormal electroencephalogram (EEG) patterns and brain lesions found in MRI images [45]. It is the fourth commonest neurological

problem after migraines, stroke and Alzheimer's disease [41]. It affects about 1% of the population worldwide [21] distributed across all ages and is more often seen in children and elderly people.

Epilepsy presents symptoms that may vary greatly, from full body convulsions to barely noticeable staring into space. Epileptic seizures are typically classified into two large groups [24]: generalized seizures and partial seizures, depending on whether both hemispheres or one hemisphere is involved at the onset of the seizures. A generalized seizure is typically easier to recognize, e.g., sudden falling from standing position, loss of bladder or bowel control, loss of consciousness and body shaking. Partial seizures, also known as focal seizures, have a more restricted and confined presentation, and the manifestations are dependent on the region of the brain that is convolved.

Epilepsy may have significant impacts on people's life. It is reported [91] that the rate of sudden unexpected death in epilepsy is two times higher than that in the general population. A large number of epilepsy patients suffer from physical injuries including cuts, bruises, burns, head trauma, broken bones, and others. People with epilepsy are more likely to manifest signs of mood disorders (such as depression and suicidal thoughts) than people without epilepsy. Low quality of sleep is another common problem in people with epilepsy. Therefore, a person's ability to live and work normally is compromised because of epilepsy.

There are several treatments available to control epileptic seizures. Antiseizure medications are one of the most effective treatments for people with epilepsy. Approximately, 70% of patients become seizure-free by taking medications regularly [39]. Surgery, by removing a small portion of brain tissue that is responsible for epilepsy, is an alternative treatment for people whose seizures can not be controlled by medications. The success rate of surgery has improved significantly after the advent of EEG and MRI, which can locate the

epileptogenic lesions accurately. Furthermore, vagus nerve stimulation (VNS) is another treatment to refractory seizures by sending electrical pulse regularly to the brain using a stimuli device. Although the above treatments are effective for most cases, the causes for epilepsy are still not well known to physicians. Some possible causes include genetic defects, brain tumors, cavernous malformation, hippocampal sclerosis, and brain infections.

Epilepsy is diagnosed based on a thorough clinical history and neurological assessment. These patients are often investigated using tests such as EEG, MRI, and other functional imaging. EEG signals, which are collected by electrodes with one end attached to the patient's scalp and the other end connected to an EEG machine, can show the patterns of brain electrical activity. Doctors can examine these brain waves or EEG signals to determine whether they are epileptogenic or not. Neuroimaging, such as computed tomography (CT) or MRI, is another test for investigating epilepsy by taking pictures of the brain so that structural abnormalities of the brain can be easily seen by doctors. MRI is usually preferred as it is able to generate high-quality images of the brain.

1.2 MRI-Based Diagnosis

The development of imaging techniques in medicine, especially the improvement of MRI, has dramatically changed the process for epilepsy diagnosis and treatment [21]. Brain MRI has become an essential tool for the investigation of epilepsy since it is able to generate high-resolution images with high soft tissue contrast, which makes it possible to accurately locate subtle epileptogenic lesions in the brain before surgical treatment of epilepsy. A general introduction to MRI is given below followed by its application in imaging epileptogenic lesions.

MRI is widely used for neuroimaging because of its superior imaging abili-

ties. Fig. 1.1 [114] (a) shows an MRI scanner, which primarily consists of the main magnet, radio frequency (RF) coil and gradient coils. The main magnet creates a strong homogeneous magnetic field B_0 that aligns the protons (within hydrogen atoms) along the direction of the field. The RF coil then burst short electromagnetic waves that knock the protons off their position. When the RF electromagnetic wave is stopped, the perturbed protons will return to their initial position. During this relaxation process, radio signals are emitted, which are detected by receivers. By using the gradient coil, spatial location information is also incorporated in the received signals. Finally, a picture is created based on the signals. Fig. 1.1 (b) shows the direction of the main magnetic field B_0, the perturbed proton magnetic moment M with its longitudinal component M_z and transverse component M_{xy}, and the direction of the RF pulse. The difference in proton density between soft tissue and hard tissue makes it possible to distinguish them in MRI images. By taking advantage of the intrinsic factors of MR relaxations, such as longitudinal relaxation (T1) and the transverse relaxation (T2), different soft tissue contrast can be obtained by MRI. Fig. 1.2 shows examples of T1-weighted (T1W) and T2-weighted (T2W) brain MRI images in different views. It is observed that gray matter appears dark and white matter appears bright in T1W images, while gray matter is bright and white matter is dark in T2W images.

As MRI can generate high contrast images of soft tissue structures, it is very suitable for imaging and capturing subtle lesions in the brain. As a consequence, MRI is becoming one of the most frequently used investigation tools in the diagnosis of epilepsy. Some of the epileptogenic lesions that can be captured by MRI are cavernous malformation (CM), focal cortical dysplasia (FCD), mesial temporal sclerosis (MTS), arteriovenous malformation (AVM), heterotopias, amygdala atrophy, congenital malformations and other genetic disorders.

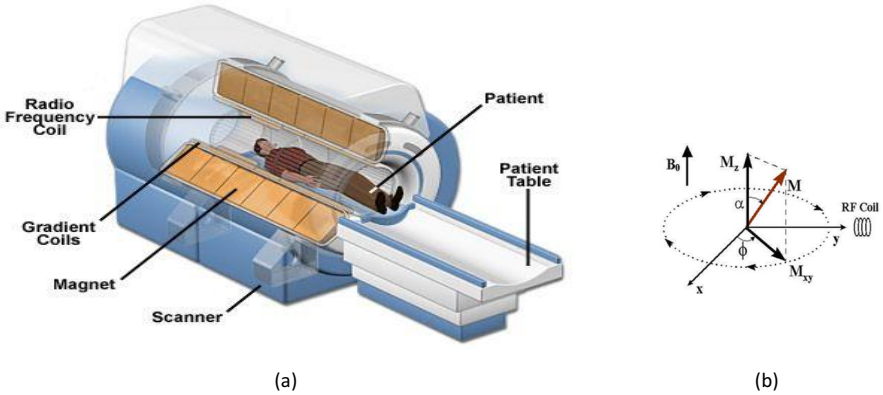


Figure 1.1: Magnetic resonance imaging. (a) MRI scanner cutaway [1], (b) a perturbed proton in magnetic field.

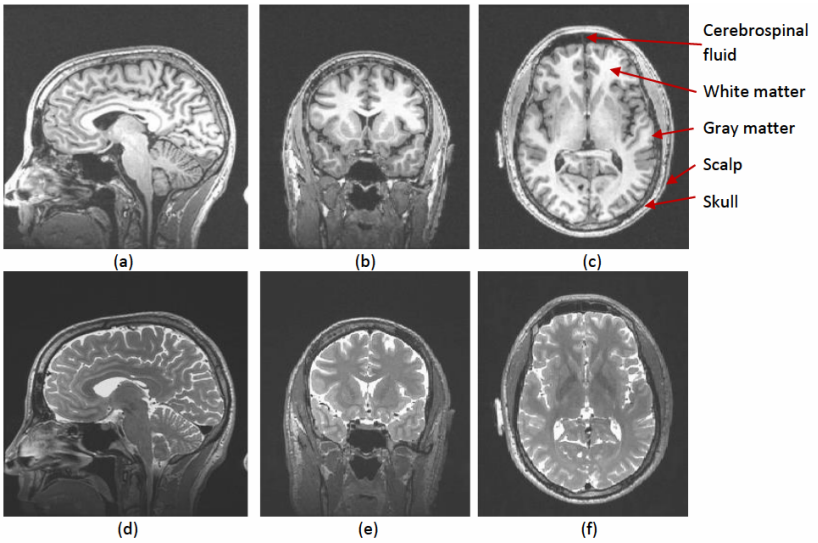


Figure 1.2: Brain MRI images. (a)(b)(c) are T1W images in sagittal, coronal, and axial view, respectively. (d)(e)(f) are T2W images in sagittal, coronal and axial view, respectively. Note that (c) is labeled with cerebrospinal fluid, white matter, gray matter, scalp, and skull.

Cavernous malformation (CM) or cavernoma is the vascular abnormality that may cause seizures, headache or vision problems. It affects a large number (about 0.5% [38]) of the general population all over the world. MRI is becoming the standard test for diagnosing cavernomas because of the high soft tissue contrast in images [21]. Fig. 1.3 shows three examples of cavernoma in T2W MRI image. In T2W images, compartments filled with water (e.g. CSF) appear bright and tissues with high-fat content (e.g. white matter) appear dark.

As cavernous malformations are clusters of abnormal thin-walled blood vessels with surrounding hemosiderin (which is caused by bleeding), they appear bright in the center with dark surroundings in T2W images.

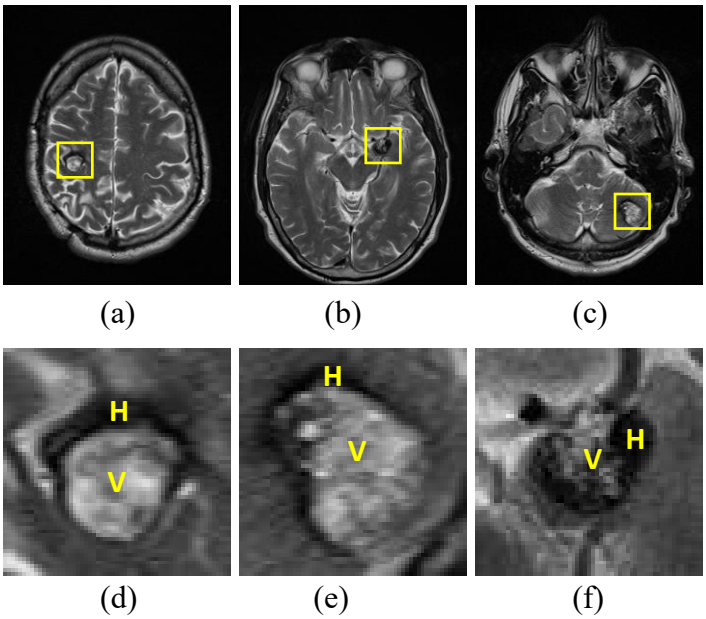


Figure 1.3: Three examples of cavernous malformation in T2W axial plane images. (a)(b)(c) are upper, middle, and lower level brain slices with cavernomas (in square boxes). (d)(e)(f) are the enlarged cavernoma regions. H: hemosiderin, V: vessels.

Mesial temporal sclerosis (MTS), also referred to as hippocampal sclerosis, is the brain lesion closely associated with temporal lobe epilepsy. It is caused by hippocampal neuronal loss accompanied by gliosis and atrophy [52], which could be captured by MRI clearly. The two primary MRI findings of MTS are hippocampal volume loss (caused by atrophy) and high intensity (related to gliosis). Fig. 1.4 shows the volume loss and high intensity of the right-side hippocampus in T2W and its fluid-attenuated inversion recovery (FLAIR) images.

Focal cortical dysplasia (FCD) is the malformation of the cortical development, which may be caused by reasons of cortical architecture or cytological abnormalities [85]. It is one of the most common causes of epilepsy in chil-

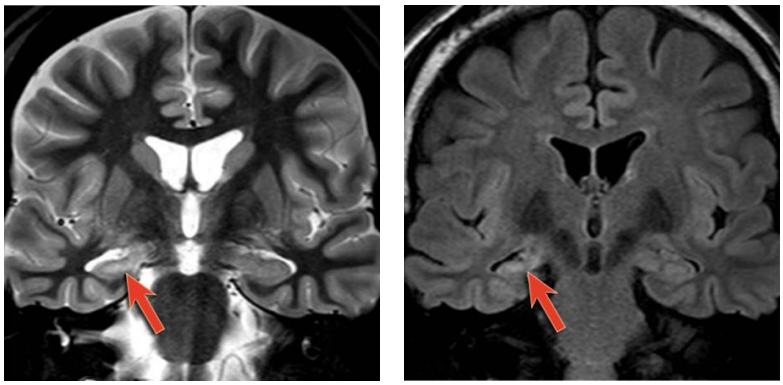


Figure 1.4: Mesial temporal sclerosis. (a) T2W image, (b) the corresponding FLAIR image.

dren [3]. MRI is the imaging methodology usually used in identifying FCD as it can provide images with high contrast and high resolution. Fig. 1.5 [125] shows two examples of FCD in T1W MRI images. A clinician typically checks MRI images for the following features to diagnosis FCD: blurred gray-white matter boundary, increased/decreased cortical thickness, and high intensity in T2/FLAIR sequence.

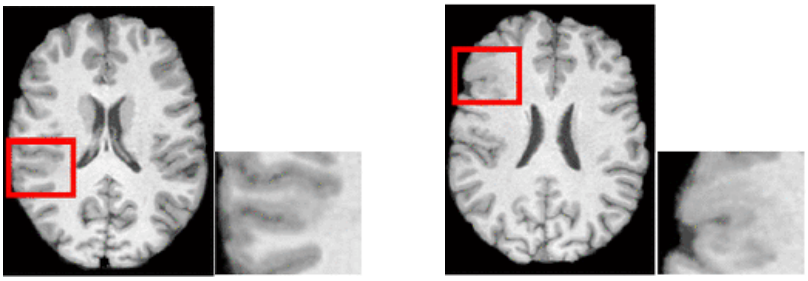


Figure 1.5: Focal cortical dysplasia in T1W images.

1.3 Problem Statements and Motivation

MRI has become a standard investigative tool for diagnosing epilepsy as it generates high-contrast and high-resolution brain images. Traditionally, neu-

rologists and radiologists visually inspect brain MRI images for the detection of epileptogenic lesions. However, inspecting MRI images for brain lesions can be a labor-intensive and subjective task even for a very experienced specialist due to the large volume of images, the subtlety of epileptogenic lesions, the complexity of brain anatomic structure and partial volume effect [45]. Therefore, it is beneficial to develop an CAD system that can detect subtle epileptogenic lesions quickly and automatically to facilitate a prompt diagnosis.

The purpose of this thesis is to develop robust and accurate CAD algorithms for the epileptogenic lesions, including CM, MTS, and FCD, based on brain MRI image analysis. Fig. 1.6 shows the overall framework of the proposed system. Before the detection of epileptogenic lesions, image preprocessing is usually performed on the input brain MRI image. Typically, it includes skull stripping to remove the non-brain structures from image, brain segmentation to obtain the gray matter, white matter and CSF tissue, brain registration to align the image with a standard brain atlas, and 3D reconstruction to generate the surface models of cortical and sub-cortical structures. The detection of epileptogenic lesions is then performed based on the image features and classification models. In this work, the automatic detection algorithms for CM, FCD, and MTS have been developed. The detection results of these lesions are finally obtained to assist doctors in the diagnosis of epilepsy. The proposed system is expected to improve the efficiency of epilepsy diagnosis and provide supportive or complementary information to physicians.

1.4 Major Contributions of the Thesis

The major contributions of this thesis are listed below:

Contribution I: A computer-aided detection technique is proposed for cavernous malformation using support vector machines and multiple texture features of MR images [152, 154, 155] (presented in Chapter 3). To the best of

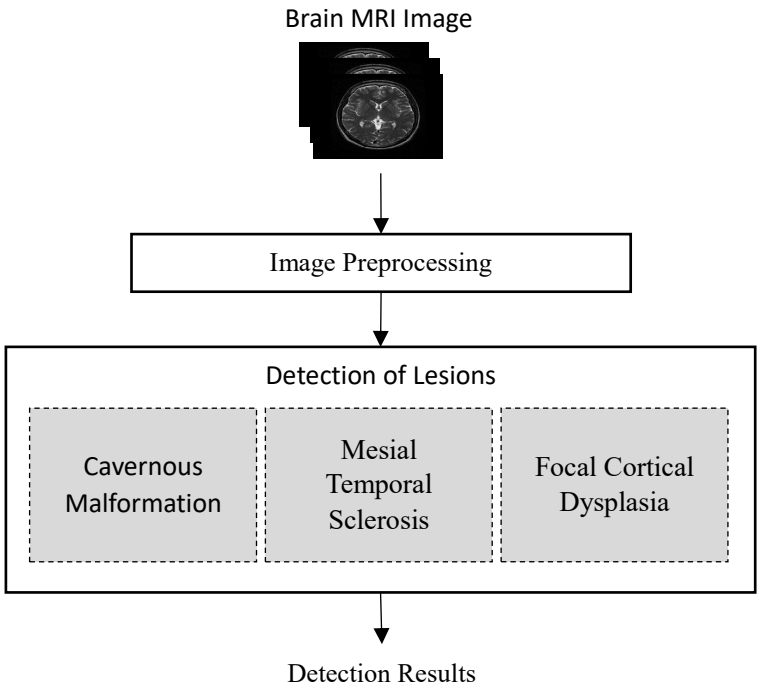


Figure 1.6: The framework of the proposed CAD system for epileptogenic lesions.

the author's knowledge, this is the first technique proposed specifically for the detection of cavernous malformation.

Contribution II: A computer-aided detection technique is proposed for mesial temporal sclerosis based on the hippocampus and cerebrospinal fluid features in 3D MR images [149, 151, 153] (presented in Chapter 4). In this technique, the hippocampal shape is modeled using the principal component analysis and the mesh model of the hippocampus. The cerebrospinal fluid feature is proposed for the detection of mesial temporal sclerosis and showed its effectiveness in experiments.

Contribution III: An automated detection technique is proposed for focal cortical dysplasia using a deep convolutional neural network [150] (presented in Chapter 5). In this technique, an automated patch extraction method is proposed based on the brain gray matter probability map and the Radon

transform. An CNN architecture is proposed to classify image patches, and it demonstrated superior performance over the classifiers using manually-crafted features.

1.5 Organization of the Thesis

The rest of this report is organized as follows: Chapter 2 presents a broad literature review on image processing techniques that are widely used in brain MRI image analysis. Chapter 3 presents the proposed technique for the automated detection of cavernous malformation based on machine learning. Chapter 4 presents the proposed technique for the detection of mesial temporal sclerosis based on 3D features and classification. Chapter 5 presents the proposed technique for the detection of focal cortical dysplasia using deep learning. The main conclusions and proposed future work are presented in Chapter 6, which is followed by the bibliography.

Chapter 2

Literature Review

The purpose of a CAD system is to process the input images and output useful information (e.g., location, volume or shape of suspected lesions) for diagnosis, which may serve as a "second opinion" to assist an expert radiologist in interpreting medical images so that the efficiency and accuracy of diagnosis could be improved [8]. Fig. 2.1 shows the general framework of a typical CAD system. It includes four major steps: preprocessing, feature extraction, and classification. In this chapter, a brief overview of some common computer-aided image analysis techniques for MRI images is presented. Section 2.1 reviews the preprocessing techniques for MRI images. Section 2.2 and section 2.3 present, respectively, the image segmentation techniques and typical image features for medical image analysis. In section 2.4, some popular classification models are reviewed. Existing CAD systems based on MRI image analysis are briefly reviewed in 2.5, followed by a summary in section 2.6.

2.1 Preprocessing

The accuracy of digital images is usually compromised by noise, intensity variation, and inhomogeneity. To solve these problems, the images are typically preprocessed by one or a combination of the following algorithms: smoothing, inhomogeneity correction, and intensity standardization. Brain MRI images

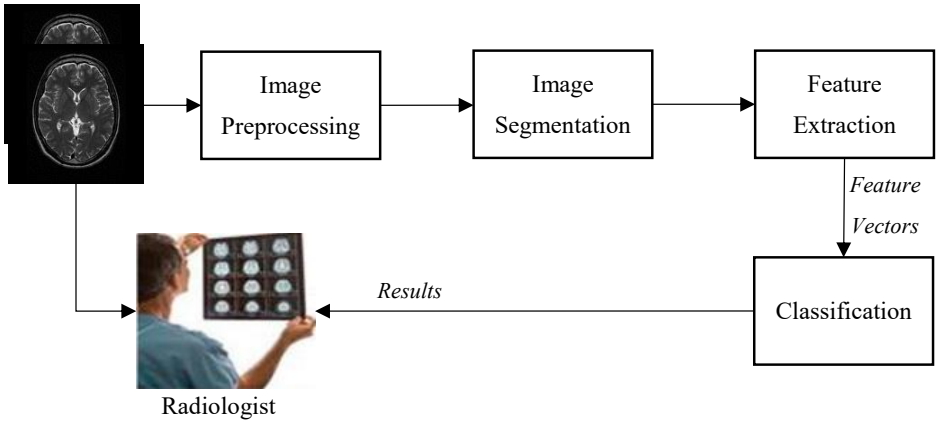


Figure 2.1: General framework of a CAD system.

also vary in size, shape, and orientation at acquisition. Therefore, registration with a standard brain atlas is often necessary in preprocessing.

2.1.1 Smoothing

Smoothing is a technique to reduce noise in an image typically by applying low pass filters to the image. Four most commonly used filters are the mean filter, median filter, Gaussian filter, and anisotropic filter. Anitha *et al.* [6] used mean filter to smooth brain MRI image in preprocessing stage. Beheshti *et al.* [16] used an 8-mm full-width-half-maximum Gaussian kernel to smooth the gray matter. Zijdenbos *et al.* [171] used anisotropic diffusion to reduce noise in MRI.

2.1.2 Inhomogeneity Correction

Due to the non-uniformity of the magnetic field and RF coil in MRI, intensity inhomogeneity is a common issue in MRI images, even though it is hardly noticeable by human eyes [147]. Fig. 2.2 shows an example of intensity inhomogeneity [147] in a brain MRI image and we can see the low-frequency spatial intensity variation is removed in the corrected image of Fig. 2.2. Meier *et al.* [110] already used the inhomogeneity correction techniques to remove

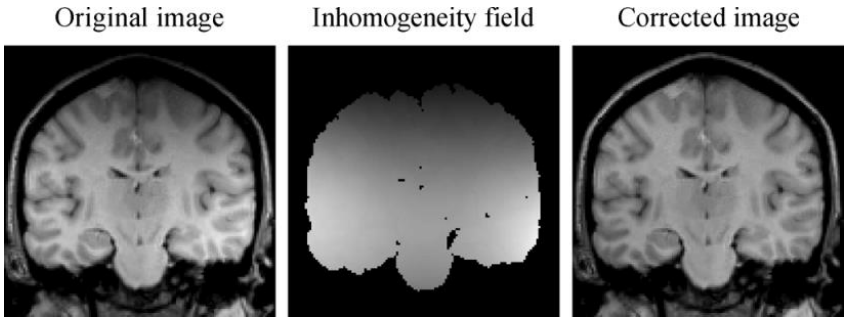


Figure 2.2: Intensity inhomogeneity in brain MRI image. Note that the top parts of the brain area have lower intensities compared to the bottom parts as shown in the inhomogeneity field.

the intensity biased field in the CAD system for the brain tumor.

2.1.3 Intensity Standardization

It has been observed that pixel intensities of a brain MRI of a patient vary over different scans even if the same scanner is used [104]. To make the CAD technique robust to intensity variation, several algorithms have been proposed for intensity standardization by manipulating the image histogram, such as histogram matching [115], histogram estimation with Gaussian mixtures [63] and joint histograms registration [78]. Zhuge *et al.* [170] suggested that brain MRI image segmentation can be simplified by using intensity standardization in the preprocessing stage.

2.1.4 Registration

The purpose of registration is to find the optimal transformation t' that is able align the input image into the same coordinate system with the reference image. In general, it can be calculated using the following equation:

$$t' = \arg\max_{t} \mu(I(\mathbf{x}), R(t(\mathbf{x}))) \tag{2.1}$$

where I is the input image, R is the reference image, μ is similarity measurement (such as mean square error and mutual information), and $t(\mathbf{x})$ is

transformation (linear or nonlinear) to the coordinate $\mathbf{x}$. On the one hand, registration can be applied to images of different modalities (ultrasound, CT, MRI, etc.) for the same subject. For example, Wein *et al.* [160] proposed a fully automatic registration technique using the affine transformation that aligns CT and ultrasound images to provide better localization for liver or kidney surgery. On the other hand, registration of the same modality between different subjects is also commonly used in medical CAD systems. Ashburner *et al.* [11] proposed a brain tissue segmentation method based on image registration with a standard probability atlas of the brain.

2.2 Segmentation

In computer-aided diagnosis systems, segmentation of regions of interest (ROI) is a typical step before any subsequent quantitative analysis, such as feature extraction and classification for lesion identification. This section reviews some basic techniques in segmentation that are widely used in skull stripping, brain tissue segmentation, and candidate region detection.

2.2.1 Thresholding

Thresholding is a simple yet powerful image segmentation method, and it uses the intensity image histogram to find thresholds that separate image into objects. Shan *et al.* proposed to use Otsu's thresholding method [117] for background and foreground separation followed by morphological operations in the skull stripping algorithm.

2.2.2 Clustering

Clustering, which is an unsupervised segmentation method, is a process of grouping pixels into different clusters, where the pixel intensity values in the same group (a cluster) are more similar to each other than those in other

groups. Anitha *et al.* [6] used the k-means clustering technique [28] to segment the brain to different clusters before feature extraction. Li *et al.* proposed to use mean-shift clustering technique [108] for corpus callosum segmentation in brain MRI images.

2.2.3 Active Contour Models

A widely used approach for object segmentation is based on active contour models (ACMs) or snakes. The basic idea of these algorithms is to evolve a continuous curve in a 2-D plane to delineate structures in an image. The evolution of the ACMs is subject to constraints from the given image by seeking to minimize an energy function based on gradient or intensity information. One widely used model is based on gradient vector flow (GVF) [164], which is calculated as follows:

$$E = \int_0^1 E_{int}(\boldsymbol{x}(s)) + E_{ext}(\boldsymbol{x}(s))ds \tag{2.2}$$

where $\boldsymbol{x}(s) = [x(s), y(s)],\ s \in (0, 1)$ is the contour. E_{int} is the internal energy calculated as follows:

$$E_{int}(\boldsymbol{x}(s)) = \frac{1}{2}[\alpha|\boldsymbol{x}'(s)|^2 + \beta|\boldsymbol{x}''(s)|^2] \tag{2.3}$$

where $\boldsymbol{x}'(s)$ and $\boldsymbol{x}''(s)$ are the first and second derivatives of $\boldsymbol{x}(s)$, respectively, and α and β are weighting parameters. E_{ext} is the external energy of the contour calculated as follows:

$$E_{ext} = \int\int \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\bigtriangledown f||\boldsymbol{v} - \bigtriangledown f|^2 dxdy \tag{2.4}$$

where $\boldsymbol{v}(x, y) = (u(x, y), v(x, y))$ is the gradient vector flow, f is the gradient of smoothed image as the edge map, $\{u_x, u_y, v_x, v_y\}$ are partial derivatives with respect to x and y. Yoon *et al.* [166] proposed another active contour model using Gaussian gradient force as the external force.

2.2.4 Atlas-Based Segmentation

Atlas-based segmentation approaches are becoming more and more popular in segmenting brain tissues and brain structures. These approaches typically align the input image with a standard atlas that includes the anatomy, morphology, and soft tissue information of the regions of interest. The segmentation can be achieved by either direct extraction of the target region or iteratively refinement within a probabilistic framework. Aljabar *et al*. [4] proposed to use multiple brain atlases in the segmentation of brain subcortical structures. Heckemann *et al*. [62] proposed an automatic brain segmentation technique based on the registration with 29 atlases and decision fusion. The most popular brain atlas is MNI152 [48] created by the International Consortium for Brain Mapping (ICBM) using a high-resolution MRI image data set. All MRI scans are first non-uniform intensity corrected and intensity normalized. These scans are then co-registered in the MNI stereotaxic space. Finally, they are average voxel-by-voxel and the brain atlas is created.

2.3 Feature Extraction

In this section, we review four frequently used image features in brain MRI image analysis. They are: (1) first-order textures, (2) histogram of oriented gradients (HOG), (3) local binary pattern (LBP), and (4) discrete wavelet transform (DWT) coefficients. First-order textures provide information about the intensity distribution, and HOG describes the edge orientation, while LBP provides the pixel correlations in a small neighboring region. The DWT is related to the frequency and time resolution properties of an image.

2.3.1 First-order Textures

In medical image processing, textural features are important in tissue classification and lesion detection as the appearance of different tissues or lesions present differently in images. The texture features include mean, standard deviation, contrast, smoothness, etc. (as listed in Table 2.1), and they are calculated based on the intensity histogram of an ROI.

Table 2.1: First-order texture features.

Name	**Expression**
Mean	$m = \sum_{i=0}^{L-1} z_i p(z_i)$
Standard deviation	$\sigma = \sqrt{\frac{1}{N-1}(\sum_{i=0}^{L-1} (z_i{}^2 h(z_i)) - Nm^2)}$
Smoothness	$s = 1 - \frac{1}{(1+\sigma^2)}$
Third moment	$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$
Uniformity	$u = \sum_{i=0}^{L-1} p^2(z_i)$
Entropy	$e = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$

where R is an image region with N pixels, z_i is a pixel intensity, $h(z)$ is the histogram of z, $p(z)$ is probability density function of z, and L is the number of the intensity levels.

2.3.2 Histogram of Oriented Gradient

The histogram of oriented gradient (HOG) feature describes the appearance of an object using the distribution of local intensity gradients and edge direction. The HOG feature vector of an image window is calculated as follows. First, the horizontal component g_h and the vertical component g_v of the gradient at a pixel are obtained by applying the filter kernels: $[-1, 0, 1]$ and $[-1, 0, 1]^T$, respectively. The orientation of the gradient is then calculated by $\theta = \arctan \frac{g_v}{g_h}$. The orientation θ is quantized into one of the nine values: $\hat{\theta} = \frac{2(k-1)\pi}{9}$, $1 < k < 9$. After that, the orientation histogram H of the image

window is calculated from $\hat{\theta}$ weighted by the gradient magnitude. Finally, the HOG feature of the image window is obtained by normalizing H using the following equation:

$$\hat{H} = \frac{H}{\|H\|} \tag{2.5}$$

where $\|.\|$ is the L^2 norm. Note that the HOG feature is a vector with 9 elements.

2.3.3 Local Binary Pattern

The local binary pattern (LBP) has been used successfully in computer vision [158]. The LBP feature of an image window is extracted by the following three steps. First, to calculate the LBP value of a pixel, the pixel value is compared with the eight neighboring pixels, and corresponding eight binary numbers are generated (as shown in Fig. 2.3). The LBP value of the pixel is the decimal equivalent of the concatenated 8-bit binary numbers. Second, the histogram of the LBP values for all pixels in the candidate region is calculated (the histogram size is 1×256). Finally, the uniform pattern [146] (by choosing the lowest decimal values of all rotationally symmetric LBP patterns) can be used to reduce the size of the feature vector to 1×59.

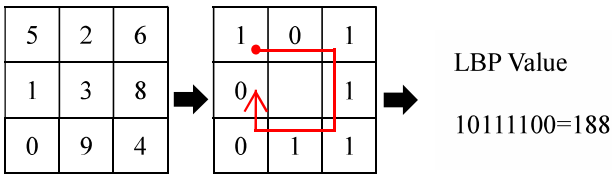


Figure 2.3: LBP value for the center pixel.

2.3.4 The Discrete Wavelet Transform

The discrete wavelet transform (DWT) is a technique to separate data into different scales and it captures both frequency and location information. Fourier

series analysis chooses sinusoids as the basis function whereas wavelet analysis is based on orthonormal basis functions. Sinusoids are useful in analyzing periodic and time-invariant phenomena, while wavelets are well suited for the analysis of transient, time-varying signals because its energy concentrated in time. Furthermore, in the spatial domain, DWT analysis also gives the best performance in detecting discontinuities or abrupt changes in signals. Anitha *et al.* [6] proposed to use discrete wavelet transform (DWT) coefficients as feature vectors for tumor classification.

2.4 Classification

After feature extraction, a candidate region can be classified into lesion or non-lesion using machine learning models. In this section, four widely used models are reviewed.

2.4.1 K-Nearest Neighbors

The k-nearest neighbors (KNN) is a simple classification model. To determine the class of a new instance, the classes of k nearest neighbors are obtained and the major class is assigned to the new instance. Sometimes the neighbors are weighted by the inverse of their distance for better classification. The best choice of K is obtained at the smallest validation error. KNN has already been successfully used in brain tumor classification [6][103] and brain abnormality segmentation [88].

2.4.2 Logistic Regression

The logistic regression (LR) is a simple but popular binary classification model. It converts the continuous value of a linear function of inputs to a probability value between 0 and 1 using the logistic function, as shown in the following

equation.

$$P(y=1|\mathbf{x}) = h_\theta(\mathbf{x})\frac{1}{1+\exp(-\theta^T\mathbf{x})} \tag{2.6}$$

where $\mathbf{x}$ is the feature vector, y is the observation and θ are the parameters that can be obtained by minimizing the following cost function $J(\theta)$.

$$J(\theta) = -\sum_i (y^{(i)} log(h_\theta(\mathbf{x}^{(i)})) + (1-y^{(i)}) log(1-h_\theta(\mathbf{x}^{(i)})) \tag{2.7}$$

where $y^{(i)}$ is the binary label (0 or 1) for the ith instances. It can be solved using the optimization algorithms such as gradient descent. Wang *et al.* [157] proposed a MS detection technique using logistic regression classifier based on DWT features.

2.4.3 Support Vector Machines

Support vector machines (SVMs) are supervised learning models popularly used in binary classification problems. In a training stage, a maximum margin hyperplane is determined, which means the distance from the hyperplane to the nearest data points of each class is maximized so that the generalization error of classification or prediction is minimized. The SVM is obtained by solving the following optimization problem:

$$\begin{aligned} &min_{\mathbf{w},b}\ \frac{1}{2} \parallel \mathbf{w} \parallel^2 \\ &s.t.\ (\mathbf{w}^T\mathbf{x}_i + b)y_i \leq 1,\ i = 1, ..., m \end{aligned} \tag{2.8}$$

where $\mathbf{w}$ is the normal vector to the hyperplane, m is the size of training set, $\mathbf{x}_i$ is the feature vector of the ith block and y_i is its label. After solving this quadratic programming problem, the separating hyperplane is obtained as $\mathbf{w}^T\mathbf{x}+b=0$. For data that is not linearly separable, SVMs can take advantage of kernel trick $k(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i)\cdot\varphi(\mathbf{x}_j)$ to transfer data to be linearly separable in higher dimension.

In the testing phase, given a feature vector $\mathbf{x}$ corresponding to an image

block, the image block is classified as class 1 if $\mathbf{w}^T\mathbf{x} + b > 0$, otherwise the image block is classified as class 2.

2.4.4 Artificial Neural Networks

Inspired by biological neural systems, the artificial neural networks (ANNs) are comprised of artificial neurons (nodes) and interconnections with weights. The architecture of an ANN includes an input layer, multiple hidden layers, and an output layer, and each layer contains many nodes. Fig. 2.4 shows an example of ANN, and it is a three-layer feed-forward network. To determine the interconnection weights, backpropagation is the most popularly used supervised learning algorithm in ANNs. Yamashita *et al.* [165] proposed to use

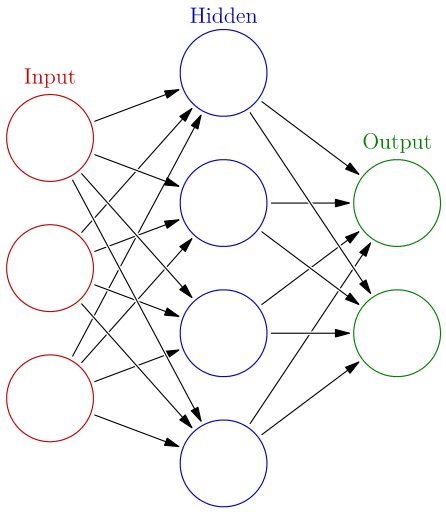


Figure 2.4: A three-layer ANN.

an ANN classifier for detecting ischemic lesions in MRI images. The ANN with 15 inputs, 8 hidden layers and 1 output was trained using the standard backpropagation algorithm. Anitha *et al.* [6] used ANN for brain tumor detection in MRI images.

2.4.5 Deep Learning

In recent years, deep learning techniques and especially convolutional neural networks (CNN) have shown great potentials in image classification problems since they could learn optimal features automatically [99]. Unlike ANN with fully connected layers, CNNs use convolutional and pooling layers to keep the number of parameters trainable and also make the network deeper. Anthimopoulos *et al*. [7] proposed a lung pattern classification technique for interstitial lung disease based on a deep CNN architecture of 5 convolutional layers with 2×2 kernels, a leaky rectifier linear unit (LReLU), and average pooling. Pereira *et al*. [121] proposed a brain tumor segmentation technique based on a CNN architecture with 4-6 convolutional layers, 3×3 kernels, and LReLU.

Some well-known deep learning architectures for image classification are AlexNet [97], VGG [136], GoogLeNet [142], and ResNet [61]. AlexNet uses rectified linear units (ReLU) as nonlinear activations, a dropout technique to avoid overfitting, and GPU to reduce training time. Instead of using large-size filters in convolutional layers, VGG networks use much smaller 3×3 filters to reduce the number of parameters. GoogLeNet further reduces the computational burden of deep networks while obtaining good performance by using the inception architecture, especially the introduction of 1×1 convolutional blocks. ResNet feeds the output of two successive convolutional layers and the input to the next layer, and bypassing two layers improves the performance of deep learning.

2.4.6 Performance Evaluation Metrics

The purpose of the evaluation is to assess the performance and robustness of a CAD system. Usually, a training database is used to obtain the optimal model parameters and a testing database is used for performance evaluation. As the medical image databases are typically small, a popular evaluation test

is k-fold cross-validation, in which the training data is randomly separated into k dataset. Of the k dataset, k-1 dataset is used as training data and the remaining one dataset is used for testing. The kind of test is repeated by k times. If k is equal to the number of cases (patients), the k-fold cross-validation is known as leave-one-out cross-validation.

To quantitative evaluate the performance of a binary classification system, statistical measures, such as sensitivity, specificity, accuracy and precision, are commonly used. They are calculated as follows:

$$Sensitivity = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100\% \tag{2.9}$$

$$Specificity = \frac{N_{TN}}{N_{TN} + N_{FP}} \times 100\% \tag{2.10}$$

$$Accuracy = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{FP} + N_{FN} + N_{TN}} \times 100\% \tag{2.11}$$

$$Precision = \frac{N_{TP}}{N_{TP} + N_{FP}} \times 100\% \tag{2.12}$$

where N_{TP}, N_{TN}, N_{FN}, and N_{FP} are the count of true positives, true negatives, false negatives and false positives, respectively. Sensitivity is also known as recall, or true positive rate. Sometimes, precision and recall are used together as a single combined measurement for the system, known as the F-measure. It is calculated as follows:

$$f = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{2.13}$$

The receiver operating characteristic (ROC) curve is a plot that illustrates the performances of a classifier. It is obtained by plotting the true positive rate ($Sensitivity$) against the false positive rate ($1 - Specificity$) at various threshold points. The ROC curve also provides a tool to find the optimal classification models.

To measure the performance of a segmentation technique, Dice coefficient is often used as the similarity measurement. If A corresponds to the automatically segmented region and B is the ground truth, the Dice coefficient is

calculated as follows:

$$d(A, B) = \frac{2|A \cap B|}{|A| + |B|} \tag{2.14}$$

where $|A|$ and $|B|$ are the areas of image region A and B, respectively, and $|A \cap B|$ is the mutual area.

2.5 CAD Systems

Table 2.2: Related CAD systems based on brain MRI image analysis. BT: brain tumor, AD: Alzheimer's disease, MS: multiple sclerosis, Dis.: disease, Cla.:classifier, Acc.: accuracy reported.

Reference	**Dis.**	**Images**	**Features**	**Cla.**	**Acc.**
Bauer *et al.*, 2011 [14]	BT	T1, T1C, T2, FLAIR	Intensity and texture	SVM	
Saha *et al.*, 2012 [131]	BT	T1C, T2			92%
Meier *et al.*, 2014 [110]	BT	T1, T1C, T2, FLAIR	Appearance and context	Decision Forest	84%
Anitha *et al.*, 2016 [6]	BT	T2	DWT	ANN, KNN	92%
Kloppel *et al.*, 2008 [92]	AD	T1	GM intensity	SVM	96%
Beheshti *et al.*, 2016 [16]	AD	T1	Intensity	SVM	92%
Zhang *et al.*, 2016 [168]	AD	T1	Morphological	SVM	84%
Zijdenbos *et al.*, 2002 [171]	MS	T1, T2, PD	Modality and tissue priors	ANN	81%
Khayati *et al.*, 2008 [89]	MS	FLAIR	Intensity	Thresholding	83%
Wang *et al.*, 2016 [157]	MS	T2	DWT	LR	98%

Bauer *et al.* [14] proposed a fully automatic voxel-wise classification technique for brain tumor segmentation using a support vector machine (SVM)

combined with hierarchical conditional random field (CRF) regularization. The intensity and first-order texture features from different modalities are used for training the classifier and making classification.

Saha *et al.* [131] proposed a fast brain tumor segmentation system using the dissimilarity measurement (Bhattacharya coefficient) of two brain hemispheres to locate candidate tumor bounding boxes in each slice. A mean-shift clustering based on the centroids of the bounding boxes is then used to segment and quantify the tumor region approximately.

Meier *et al.* [110] proposed a brain tumor segmentation method using MRI images of multiple modalities (T1, T1C, T2, and FLAIR). After preprocessing (smoothing, bias-field correction, and intensity normalization), the appearance-sensitive (intensity values, first-order texture, and gradient) and context-sensitive (atlas-normalized coordinates and ray feature) features are extracted. A voxel-wise classification based on decision forest is then applied refined by a spatial regularization.

Anitha *et al.* [6] proposed a CAD system for brain tumor using a two-tier classification technique. First, the brain MRI is preprocessed using a mean filter to remove noise and morphology-based skull stripping technique to extract the brain portion. Second, a K-means clustering method is used for brain segmentation. After that, discrete wavelet transform coefficients are used to calculate feature vectors for classification. Finally, the self-organizing map neural network and K-nearest neighbor classification algorithm are used to classify brain images into cancer or non-cancer images.

Kloppel *et al.* [92] proposed to use SVM to classify T1W GM from AD patients. In this method, brain images are segmented into gray matter, white matter and cerebrospinal fluid using statistical parametric mapping (SPM) [161].

Beheshti *et al.* [16] proposed a CAD system for Alzheimer's disease (AD)

using 3D T1W brain images. First, voxel-based morphometry is applied to gray matter to obtain the volume of interests (VOIs). After extracting the intensity value of VOIs as the raw feature vector. Seven feature ranking methods are used to obtain the optimized feature vector that can minimize the classification error. Finally, the optimized feature vector and the SVM classifier are used for the detection of the AD.

Zhang *et al.* [168] proposed a three-stage CAD system for AD. First, AD landmark model is generated by comparing local morphological features after linear and non-linear registration. Second, a regression-forest based landmark detection method is used to identify AD landmarks in a testing image. Finally, an SVM classifier is used to diagnosis AD based on the morphological features around the detected landmarks.

Zijdenbos *et al.* [171] proposed an voxel-wise classification framework for MS lesions. Images are preprocessed using intensity normalization, smoothing and bias correction techniques. Registrations are performed before resampling the data into 1 mm resolution. Based six features (three modalities and three spatial tissue priors), an artificial neural network classifier is used for a voxel wised classification.

Khayati *et al.* [89] proposed an automated method to detect MS in brain FLAIR images using an adaptive mixture method and a Markov random field model. Each voxel is modeled as a combination of normal and abnormal parts, and it can be further classified into chronic, early and recent acute lesions by optimal thresholding.

Wang *et al.* [157] proposed a CAD system for multiple sclerosis (MS) based on biorthogonal wavelet transform, RBF kernel principal component analysis (RKPCA) and logistic regression. The biorthogonal wavelet transform of the input image is first calculated and the features are extracted. The RKPCA is then used to reduce the size of the feature vector. Finally, a linear regression

method is used to classify the image as MS or non-MS image.

2.6 Summary

In this chapter, we presented a brief review of the automated image analysis techniques including preprocessing, segmentation, feature extraction, and classification. Several CAD systems for different brain diseases (*e.g.* brain tumor, Alzheimer's disease, and multiple sclerosis) were also reviewed.

Chapter 3

Automated Detection of Cavernous Malformation

Cavernous malformation or cavernoma is one of the most common epileptogenic lesions. It is a type of brain vessel abnormality that can cause serious symptoms such as seizures, intracerebral hemorrhage, and various neurological disorders. Manual detection of cavernomas by physicians in a large set of brain MRI slices is a time-consuming and labor-intensive task and often delays diagnosis. In this chapter, we proposed an automated detection technique for cavernomas based on T2-weighted axial plane MRI image analysis. The proposed technique first extracts the brain area based on atlas registration and active contour model, and then performs template matching to obtain candidate cavernoma regions. Texture, the histogram of oriented gradients and local binary pattern features of each candidate region are calculated, and principal component analysis is applied to reduce the feature dimensionality. Support vector machines (SVMs) are finally used to classify each region into cavernoma or non-cavernoma so that most of the false positives (obtained by template matching) are eliminated. The performance of the proposed technique is evaluated and the experimental results show a superior performance in cavernoma detection compared to existing techniques.

3.1 Introduction

Epilepsy is the fourth most common neurological problem (after migraines, strokes, and Alzheimer's disease) [45], which affects about 1% of the population worldwide [21]. Epilepsy may have significant impacts on people's lives as it is reported [91] that the rate of sudden unexpected death in epilepsy is two times higher than that in the general population. A large number of epilepsy patients suffer from physical injuries including cuts, bruises, burns, head injury, broken bones and others. Furthermore, people with epilepsy are more likely to have mood (such as depression) and sleeping problems. Although most of the causes of epilepsy are still unknown, some epileptogenic lesions are typically found in images obtained by brain magnetic resonance imaging (MRI). An important epileptogenic lesion is cavernous malformation or cavernoma that arises due to vascular abnormality involving a cluster of abnormal vessels in brain [124]. MRI is a frequently used imaging method to diagnose cavernomas in hospitals because of the high image quality of soft tissues in the brain. Fig. 3.1 shows three T2-weighted (T2W) image with cavernous malformations surrounded by square boxes. In a T2W image, a cavernoma typically looks like popcorn with black surrounding due to the presence of hemosiderin. In typical brain lesion detection, the task is to detect the presence of a malformation and determine its location [15]. However, visual identification of cavernomas in a large set of MRI slices is a tedious and time-consuming task that may result in a slow diagnosis and even misdiagnosis. Therefore, computer-aided detection (CAD) that can process the input MRI images using a computer to generate detection results is highly desirable to improve the efficiency and accuracy of diagnosis.

To the author's knowledge, there is no CAD technique in the literature specifically for the detection of cavernous malformations. However, several

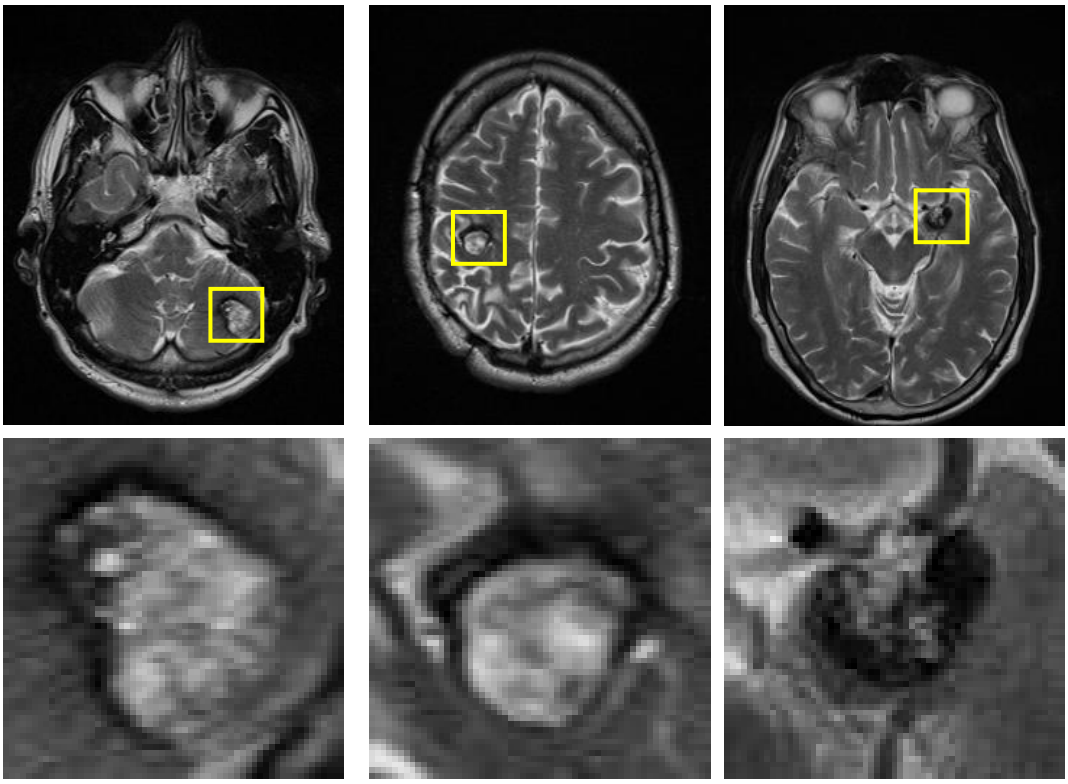


Figure 3.1: Three examples of cavernous malformations in T2W images. Top row: original images with cavernomas in yellow boxes. Bottom row: enlarged cavernoma regions.

CAD systems for other brain diseases have been proposed by researchers, such as those for brain tumors [5, 6, 43, 131], Alzheimer's disease [16, 30, 92], vascular dementia and multiple sclerosis [89, 157, 171]. Anitha *et al.* [6] proposed a CAD system for brain tumor using a two-tier classification technique. The brain MRI is first preprocessed to remove noise and skull area followed by a K-means clustering method for brain segmentation. Discrete wavelet transform coefficients are then used to calculate the feature vectors for classification. Finally, the self-organizing map neural network and k-nearest neighbor classification are used for the brain tumor classification. Saha *et al.* [131] proposed a brain tumor diagnosis system using the dissimilarity measurement of two brain hemispheres to locate candidate tumor bounding boxes, and mean shift clustering to segment and quantify the tumor. Ambrosini *et al.* [5] proposed

a template matching technique to detect metastatic brain tumors by using 3D sphere templates, and thresholding normalized cross-correlation between templates and candidate regions. Farjam *et al.* [43] proposed a method similar to [5] for detecting brain tumor but used the size, shape and brightness information of lesions to further improve the detection performance.

Beheshti *et al.* [16] proposed a CAD system for Alzheimer's disease (AD) using 3D T1-weighted (T1W) brain images. First, voxel-based morphometry is applied to gray matter to obtain the volume of interest (VOI). After extracting the intensity value of VOI as the raw feature vectors, seven feature ranking methods are used to obtain the optimized features that can minimize the classification error. Finally, the optimized feature vector and the SVM classifier are used for the detection of AD. Kloppel *et al.* [92] also proposed an AD detection technique using SVM to classify T1W gray matter segments. In this method, brain images are segmented into gray matter, white matter and cerebrospinal fluid using statistical parametric mapping (SPM) [161]. Colliot *et al.* [30] proposed an automated classification method for AD, mild cognitive impairment, and aging based on the difference of hippocampal volumes in T1W brain MRI images.

Wang *et al.* [157] proposed a CAD system for multiple sclerosis (MS) using biorthogonal wavelet transform to extract features, RBF kernel principal component analysis to reduce the size of the feature vector and logistic regression to classify the image as MS or non-MS. Zijdenbos *et al.* [171] proposed a voxel-wise classification framework for MS lesions using a three-layer artificial neural network classifier based on multi-model intensity and spatial priors. Khayati *et al.* [89] proposed an automated method to detect MS in brain FLAIR images using an adaptive mixture method and a Markov random field model. Each voxel is modeled as a combination of normal and abnormal parts, and it can be further classified into chronic, early and recent acute lesions by

optimal thresholding.

In this chapter, we propose a CAD technique to detect slices with cavernous regions in T2W brain MRI images. It is an extension of our previous work [156], in which we proposed to use skull stripping and template matching to detect cavernous malformation. This paper improves [156] in two aspects: (1) An atlas-based skull stripping technique is used instead of the morphology-based method to obtain better brain area extraction; (2) After template matching, SVM models are used to classify candidate regions to cavernomas or non-cavernomas based on texture, histogram of gradient and local binary pattern features with principal component analysis, instead of using size, shape and brightness information of candidates for classification in [156]. The remainder of this chapter is organized as follows: Section 3.2 describes the materials; Section 3.3 explains the proposed method; experimental results are presented in Section 3.4, while Section 3.5 concludes the paper.

3.2 Materials

In this study, a database of 192 T2W MRI axial images from 9 patients with cavernous malformations are used. The images were obtained at the University of Alberta Hospital. The T2W images were acquired using the turbo spin echo sequence on a 1.5T Siemens MRI scanner (TE/TR = 99/5000 ms) and have a pixel resolution of approximately 0.45×0.45mm with a slice thickness of 6-8 mm. The resolution of each image slice is 512×416 pixels. Overall, there are 23 cavernous images (out of 192), each image with one cavernoma region. Due to the large slice thickness of the T2W images in our database, the proposed detection is performed on 2D slices. Although a small data set is used in this study, the cavernomas are representative in size, location, and shape [13, 107, 145]. The cavernomas can be modeled by circular or elliptical shapes. The diameter (major/minor axes length in case of elliptical shape)

of these lesions on 2D slices spreads from 8 mm to 20 mm. Out of the 23 cavernomas, 8 are located in the frontal lobe, 8 in the temporal lobe, 6 in the parietal lobe and 1 in the occipital lobe. The shape of these cavernomas also covers a large variety. Fig. 3.2 shows six samples of cavernous malformations in our study of different size and shape. The locations and boundaries of these cavernomas are marked by an experienced neurologist and this information is used as ground truth.

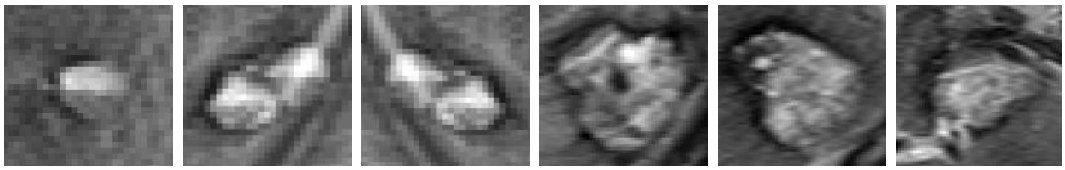


Figure 3.2: Six samples of cavernous malformations.

The T2W ICBM 2009a non-linear symmetric brain atlas [48] and its brain probability map are used in this work. The atlas, which is generated by averaging the MNI152 database, has a high resolution of 1×1×1mm and includes 189 axial slices with the size of 197×233 pixels. It is one of the most popular brain atlases in medical image analysis. The 3D atlas has a corresponding brain probability map (B) that indicates if an MRI pixel belongs to the brain area. It is defined as follows:

$$B(\mathbf{x}) = \begin{cases} 1, \ if \ \mathbf{x} \ is \ a \ brain \ pixel \\ 0, \ if \ \mathbf{x} \ is \ not \ a \ brain \ pixel \\ p, \ others \end{cases} \tag{3.1}$$

where $\mathbf{x}$ is a 3D pixel location and p is the probability that the pixel at $\mathbf{x}$ is a brain pixel.

3.3 Proposed Method

The purpose of a CAD system is to process the input images and generate useful information (e.g., location, volume or shape of suspected lesions) for diagnosis, which can assist physicians in interpreting the medical images so that

the efficiency and accuracy of diagnosis could be improved [8]. Fig. 3.3 shows the schematic of the proposed CAD system for detecting cavernous malformation. It includes four major modules: skull stripping, template matching, feature extraction, and classification. The details of these modules are presented in the following sections.

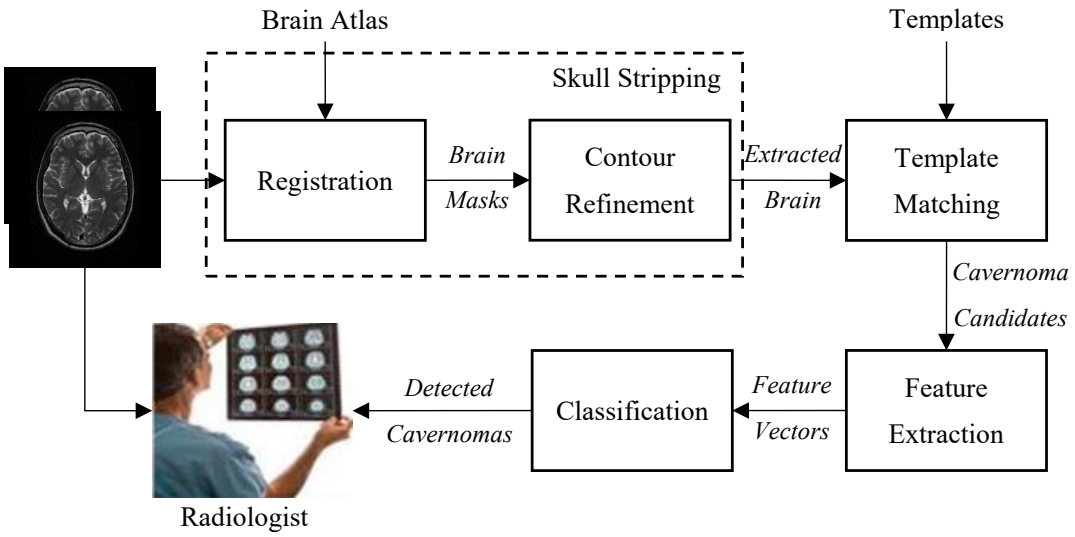


Figure 3.3: Schematic of the proposed CAD system for cavernous malformations.

3.3.1 Skull Stripping

Skull stripping, or brain extraction, is a technique that removes the non-brain tissues, such as the skin of the scalp, bone of the skull, dura matter and eyeballs, from a brain image. It is the first step in the proposed cavernomas detection system as only brain area is the region of interest (ROI) for cavernomas detection. A number of automatic skull-stripping methods have already been proposed in recent years, which can be categorized into morphology-based, intensity-based, deformable surface-based, standard atlas-based and hybrid methods [86]. The BET [137], BSE [134], and 3dSkullStrip [130] methods are widely used for skull stripping in MRI images. Among them, BSE and 3dSkullStrip are designed to work primarily with T1W MR images. The BET

method is capable of processing T2W images based on a deformable model driven by local adaptive forces. This technique does not perform well for large slice thicknesses as in our case. Therefore, we propose an improved skull stripping method using atlas registration. This method has two steps: (1) affine registration with a brain atlas to obtain the initial brain masks, and (2) contour refinement in each slice using active contour model for brain extraction.

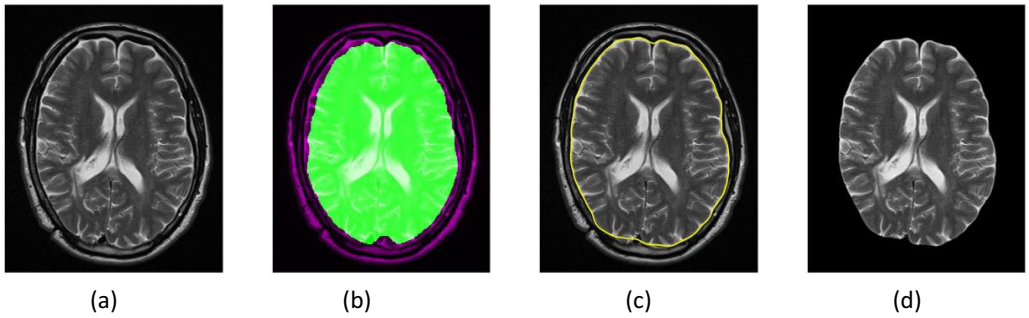


Figure 3.4: Skull stripping results of one slice of a 3D image. (a) An input T2W slice, (b) input slice overlaid with the brain mask (in green) after registration, (c) refined brain contour (shown by the yellow boundary), (d) extracted brain.

Registration

In order to obtain an approximate brain area, researchers sometimes perform registration with a standard brain atlas to generate the mask of ROI (i.e., the brain). In this work, the T2W ICBM 2009a non-linear symmetric brain atlas is used to register with the input T2W brain images to obtain the initial brain masks for contour refinement.

The brain atlas is used as the moving image, and the input image as the fixed image in a 3D affine registration framework. The mutual information of the moving and fixed image serves as the objective function to be maximized. The purpose of registration is to find the best transformation t' that will align the input image to the brain atlas in the same coordinate system. In this work,

t' is obtained by solving the following equation.

$$t' = \arg\max_{t} \mu(I(\mathbf{x}), A(t(\mathbf{x}))) \tag{3.2}$$

where I is the input 3D MRI image (about 20 axial slices), A is the reference brain atlas image (189 axial slices), μ is the mutual information, $t(\mathbf{x})$ is affine transformation to the coordinate $\mathbf{x}$, and t' is the best transformation. Note that affine transformation is a combination of transforms including translation, scale, shear and rotation. For the 3D affine transformation, the degree of freedom is 12. The mutual information of two images I and A is calculated as follows:

$$\mu(I, A) = e(I) + e(A) - e(I, A) \tag{3.3}$$

where $e(I)$ and $e(A)$ are the entropies of images I and A, respectively, and $e(I, A)$ is the joint entropy. The entropies are calculated based on the 1D probability density function of the pixel intensities of the entire 3D images I and A, and the joint entropy is calculated based on the joint histogram of I and A. As the brain atlas A typically has more axial slices but less axial plane resolution than I, both sampling and interpolation operations are applied to A (to make it of the same size as that of I) before calculating the mutual information.

After registering the brain atlas (A) with the input brain image I, we obtain the best transformation t'. As only the brain area is the ROI, we apply t' to the brain probability map (B) to generate the brain mask for the input image as follows:

$$M(t'(\mathbf{x})) = \begin{cases} 1, \ if \ B(\mathbf{x}) = 1 \\ 0, \ others \end{cases} \tag{3.4}$$

where M is the brain mask of input image I and $\mathbf{x}$ is a pixel location in B. Fig. 3.4 (a) shows one slice of a T2W brain MRI image and Fig. 3.4 (b) shows the image slice in (a) overlaid with its brain mask obtained after registration.

We observe that, even though the brain mask is not perfect, it is close enough to be an initial brain contour for further refinement.

Contour Refinement

In contour refinement, we use an active contour model, also known as snakes, which is a popular segmentation technique that minimizes the energy along the contour of object. The refinement of contours is driven by two forces: an internal force that keeps the contour smooth, and an external force that pushes it towards the object contour. The active contour model based on a gradient vector flow [164] is used in the proposed system for brain contour refinement. The contour $\boldsymbol{x}(s)$ is obtained by minimizing E as defined below.

$$E = \int_0^1 [E_{int}(\boldsymbol{x}(s)) + E_{ext}(\boldsymbol{x}(s))]ds \tag{3.5}$$

where $\boldsymbol{x}(s) = [x(s), y(s)],\ s \in (0, 1)$. The internal energy E_{int} is calculated as follows:

$$E_{int}(\boldsymbol{x}(s)) = \frac{1}{2}[\alpha|\boldsymbol{x}'(s)|^2 + \beta|\boldsymbol{x}''(s)|^2] \tag{3.6}$$

where $\boldsymbol{x}'(s)$ and $\boldsymbol{x}''(s)$ are the first and second derivatives of $\boldsymbol{x}(s)$, respectively, and α and β are weighting parameters. E_{ext} is the external energy of the contour calculated as follows:

$$E_{ext} = \iint [\gamma(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\bigtriangledown f||\boldsymbol{\kappa} - \bigtriangledown f|^2]dxdy \tag{3.7}$$

where $\boldsymbol{\kappa}(x, y) = (u(x, y), v(x, y))$ is the gradient vector flow, $\bigtriangledown f$ is the gradient of the smoothed image as the edge map, $\{u_x, u_y, v_x, v_y\}$ are the partial derivatives with respect to x and y.

Fig. 3.4 (c) shows the final brain contour (the yellow boundary) after contour refinement and Fig. 3.4 (d) is the extracted brain.

3.3.2 Template Matching

Template matching (TM) is a technique to locate regions of an image that have similar structure or pattern with a reference image known as a template. In order to increase the sensitivity of matching, multiple templates are generally used, which are transformations (scaling, rotation, etc.) of a reference template. In T2W MR images, a cavernoma appears bright at center with a dark surrounding, and it is always located inside of brain tissues that appear gray in images. Therefore, we can use templates of similar structures to match and detect the cavernous malformations. In the proposed technique, we generate the template database using the generalized Laplacian of Gaussian (gLoG) [94]. The gLoG function is given as follows:

$$\bigtriangledown^2 G(x,y) = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} \tag{3.8}$$

where (x, y) are pixel coordinates in image, and G is the Gaussian function. The definitions of G and the second derivatives of G are given as follows:

$$G(x,y) = e^{-(ax^2+2bxy+cy^2)} \tag{3.9}$$

$$\frac{\partial^2 G}{\partial x^2} = [(2ax+2by)^2 - 2a]G(x,y) \tag{3.10}$$

$$\frac{\partial^2 G}{\partial y^2} = [(2bx+2cy)^2 - 2c]G(x,y) \tag{3.11}$$

where a, b, c are parameters that control the shape and orientation of a template, and are defined as follows:

$$a = \frac{\cos^2\theta}{2\sigma_x^2} + \frac{\sin^2\theta}{2\sigma_y^2} \tag{3.12}$$

$$b = -\frac{\sin 2\theta}{4\sigma_x^2} + \frac{\sin 2\theta}{4\sigma_y^2} \tag{3.13}$$

$$c = \frac{\sin^2\theta}{2\sigma_x^2} + \frac{\cos^2\theta}{2\sigma_y^2} \tag{3.14}$$

where σ_x, σ_y are scale parameters and θ is the orientation parameter. Fig. 3.5 shows the 28 templates used in the proposed CAD system for cavernomas.

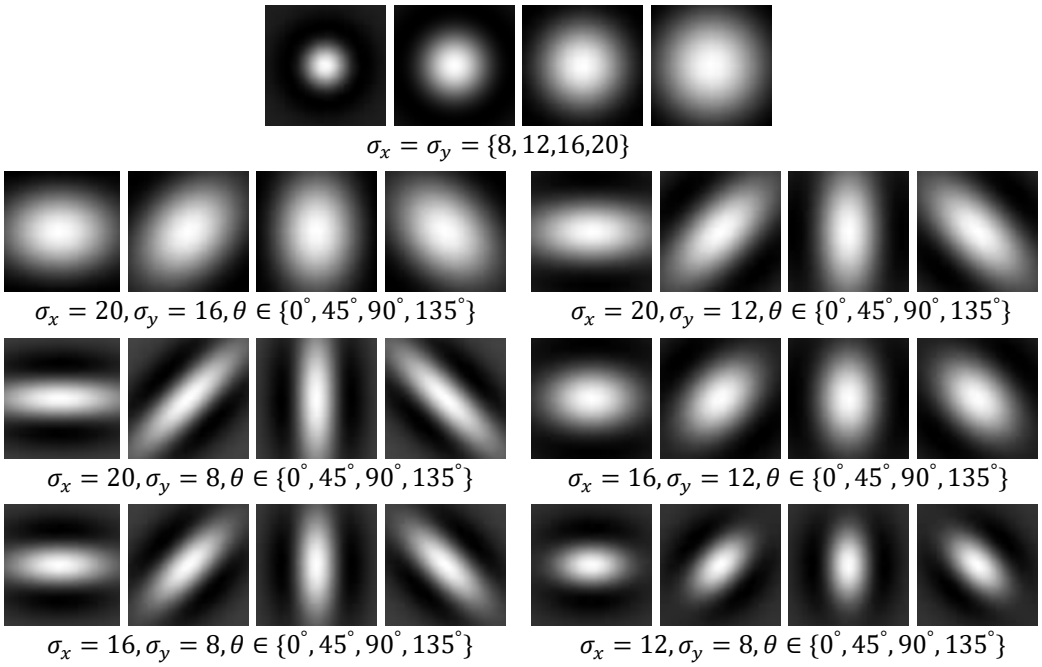


Figure 3.5: Templates used in the proposed technique.

Each template size is 25×25mm (48×48 in pixels) as the diameter of cavernous malformation varies from 8 to 20 mm in our image database. Both σ_x and σ_y are chosen from {8, 12, 16, 20} mm and θ are from {0°, 45°, 90°, 135°}.

In the proposed technique, the normalized cross correlation (NCC), defined below, is used to measure the similarity of an image region R and a template T.

$$NCC(R,T) = \frac{\sum_{i,j}[R(i,j) - \bar{R}][T(i,j) - \bar{T}]}{\sqrt{\sum_{i,j}[R(i,j) - \bar{R}]^2 \sum_{i,j}[T(i,j) - \bar{T}]^2}} \tag{3.15}$$

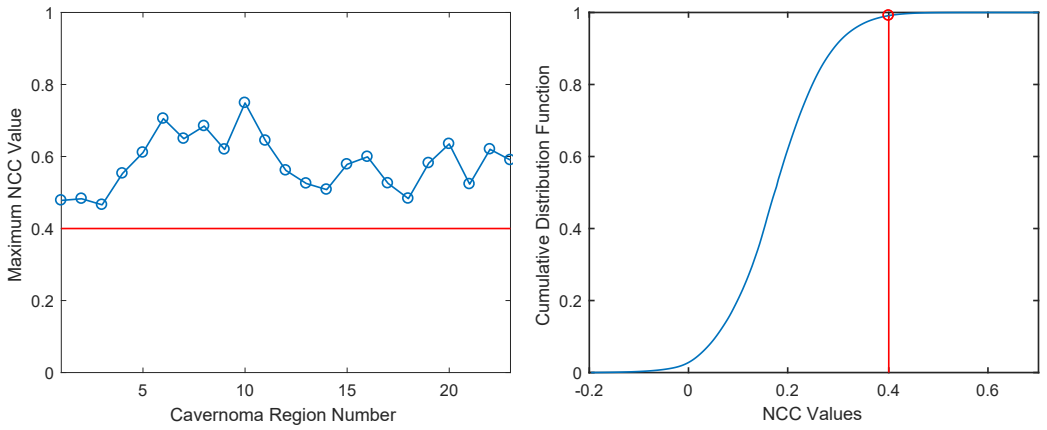
where $R(i,j)$ and $T(i,j)$ are pixel intensities of the image region and the template, $\bar{R}$ and $\bar{T}$ are the mean intensities of the region R and the template T, respectively. The NCC value is in the range of [-1, 1] and a larger NCC value reflects a better matching. To identify the matching regions, a template image is compared with the source image by sliding it (i.e., moving the template one pixel at a time from left to right and top to bottom) and a NCC value is calculated at each (central) pixel location. Using this procedure, a NCC

response map is generated for that template. As there are 28 templates in the database, 28 NCC response maps are generated for each image slice. For each response map, we find cavernoma candidates by thresholding the local maxima.

A region R is a cavernoma candidate if

$$NCC(R, T) > \tau \tag{3.16}$$

where τ is a predetermined threshold. Fig. 3.6 (a) shows the NCC values of cavernoma regions, whereas Fig. 3.6 (b) shows the cumulative distribution function (CDF) of NCC values of both cavernoma and non-cavernoma regions. In this paper, we have chosen $\tau = 0.4$ to detect candidate cavernoma regions with a low false positive rate.



(a) (b)

Figure 3.6: Distribution of NCC values. (a) Maximum NCC values for each cavernoma, (b) the plot of the NCC cumulative distribution function.

3.3.3 Feature Extraction

Using TM, we obtain a large set of cavernoma candidate regions. However, many of these regions are false positives. The feature extraction and classification modules reduce the number of false positives by classifying the region into either cavernoma (CM) or non-cavernoma (non-CM). Three kinds of features

are used: (1) first-order textures, (2) histogram of oriented gradients (HOG) and (3) local binary pattern (LBP). First-order textures provide information about the intensity distribution, and HOG describes the edge orientation, while LBP provides the pixel correlations in a small neighborhood. The three types of features together are able to capture the appearance information of an object efficiently, and effectively.

First-order Textures

In this work, we use six first-order texture features (mean, standard deviation, smoothness, third moment, uniformity and entropy), as defined in Table 2.1, based on the intensity histogram of a region. Let the texture feature vector be denoted by $\boldsymbol{\phi}$.

HOG

The HOG feature describes the appearance of an object using the distribution of local intensity gradients and edge direction. The HOG feature vector of an image region is calculated as follows.

Step 1. A candidate image region is divided into overlapping blocks and each block is comprised of cells, as shown in Fig. 3.7. The size of regions, blocks and cells are 48×48, 16×16 and 8×8 pixels, respectively. The overlap of adjacent blocks is half the block size. Overall, there are 25 blocks in a candidate region.

Step 2. The gradient of each pixel is obtained by applying two filter kernels: $[-1, 0, 1]$ and $[-1, 0, 1]^T$ to the image region. Each gradient is then quantized into one of the nine major orientations: $\frac{2(k-1)\pi}{9}$, $1 \leq k \leq 9$, and weighted by the gradient magnitude. After this, the orientation histogram H_c of each cell is obtained, and the size is 1×9.

Step 3. The feature vector of each block is then obtained as H_b= $[H_{c1}, H_{c2}, H_{c3}, H_{c4}]$, where H_{cm} is orientation histogram of the mth cell. The feature vector H_b is then normalized using the following equation.

$$\hat{H_b} = \frac{H_b}{\|H_b\|} \tag{3.17}$$

where $\|.\|$ is the L^2 norm.

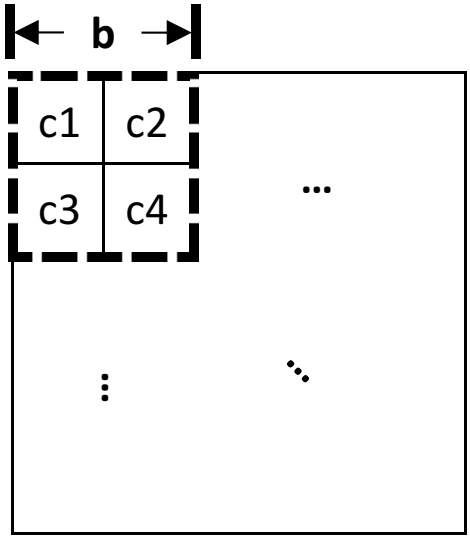


Figure 3.7: The cell and block division in an image window. b and c refer to block and cell, respectively.

The HOG feature vector of an image region is obtained by concatenating all the normalized block feature vectors. As there are 25 blocks in an image region and the size of each normalized block feature vector is 1×36, the size of the HOG feature vector of one candidate region is 1×900. Let the HOG feature vector be denoted by $\boldsymbol{\eta}$.

LBP

The LBP has been used successfully in many computer vision applications [158]. In the proposed CAD system, we calculate the LBP feature of each cavernoma candidate region. First, to calculate the LBP value of a pixel, the pixel value is compared with the eight neighboring pixels, and corresponding

eight binary numbers are generated (as shown in Fig. 3.8). The LBP value of the pixel is the decimal equivalent of the concatenated 8-bit binary numbers. Second, the histogram of the LBP values for all pixels in the candidate region is calculated (the histogram size is 1×256). Finally, we use the uniform pattern [146] (by choosing the lowest decimal values of all rotationally symmetric LBP patterns) to reduce the size of the feature vector to 1×59. Let this feature vector be denoted by $\boldsymbol{\zeta}$.

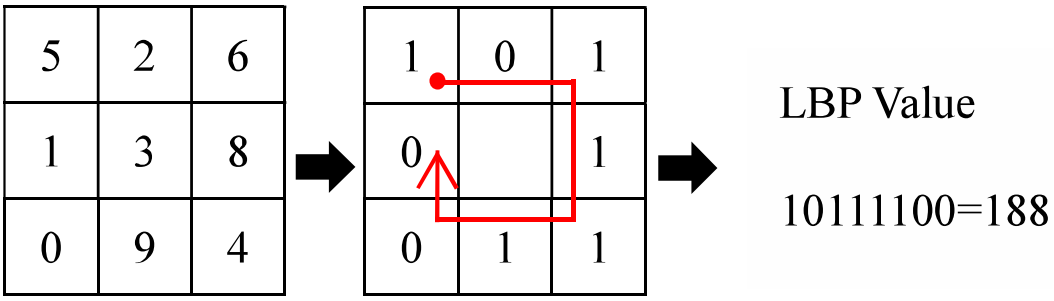


Figure 3.8: LBP value for the the center pixel.

Combining the three types of features, the overall feature vector, of dimension 1×965, is obtained as $\boldsymbol{x} = [\boldsymbol{\phi}, \boldsymbol{\eta}, \boldsymbol{\zeta}]$.

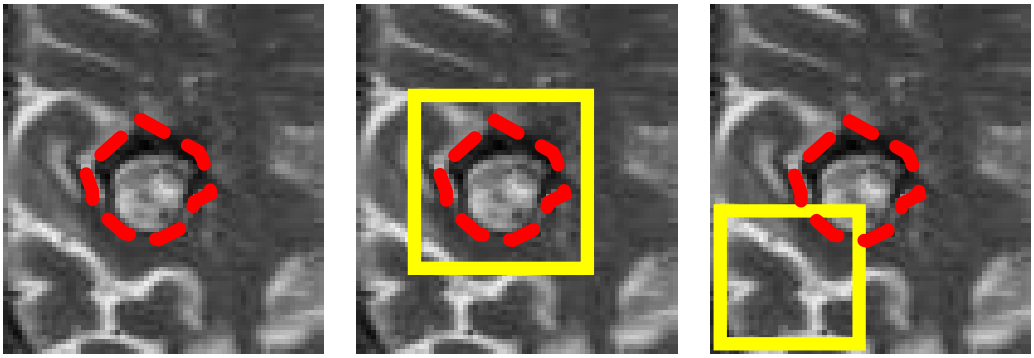
Feature Dimensionality Reduction

The feature vector $\boldsymbol{x}$ has a large dimension of 1×965. In order to discard features that are correlated with others and improve the efficiency of classification, the principal component analysis (PCA) technique is applied in our system. The PCA can reduce the feature dimension, and at the same time retain most of the variation present in the original data set [83]. In the proposed system, the number of principal components is chosen such that they can explain at least 90% of the data variance.

3.3.4 Classification

SVMs [32, 123] are widely used in classification applications from character recognition to tumor detection for its high performance and robustness even

in noisy datasets. After template matching, we typically obtain a large set of candidate cavernous regions. In order to train an SVM model, these candidate regions are labeled (by visual inspection) as either CM (1) or non-CM (-1). As shown in Fig. 3.9, a candidate region is labeled as CM if it encloses the center of a cavernous malformation, otherwise it is labeled as non-CM.



(a) (b) (c)

Figure 3.9: Candidate region labeling. (a) The ground truth boundary (red) of a cavernoma, (b) a candidate region (yellow) is labeled as CM, (c) a candidate region (yellow) is labeled as non-CM.

A common problem in medical image classification is the imbalanced dataset. The size of the two groups is highly imbalanced as there are many (about 50 times) more non-cavernoma candidates than cavernoma candidates. If we use the imbalanced dataset in the training process directly, the trained SVM tends to classify all candidates into non-CM group to obtain a higher accuracy, resulting in low sensitivity. To avoid this, we use the following methods.

Imbalanced Dataset Handling

In our CAD system, we use two techniques to avoid imbalanced training: oversampling and dataset division [95].

- Oversampling: it is a way to increase the number of cavernoma instances so that the two groups are more evenly balanced. For each cavernoma

candidate, we generate 8 more instances by sampling regions centered at the neighboring pixels of the original center pixel.

- Dataset division: although oversampling increases the number of cavernoma instances, it is still unable to balance the two groups. Therefore, we divide the dataset into multiple (e.g. 5) subsets and in each subset, all cavernoma instances and the same size of non-cavernoma instances are included. With each balanced subset, we train an SVM model. So five SVMs are trained and used for classification.

SVM

To train an SVM model for binary classification, feature vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ and their labels $\{y_1, y_2, \ldots, y_n\}$ are generated from training ground truth data, where n is the size of training set. In our CAD system, the size of each feature vector is 1×965 and each label is either CM (1) or non-CM(-1). The SVM generates a hyperplane that can separate the training data with maximum margin. The SVM model is obtained by solving the following equation:

$$\begin{aligned} &min_{\mathbf{w},b} \ \frac{1}{2} \parallel \mathbf{w} \parallel^2 \\ &s.t. \ (\mathbf{w}^T\mathbf{x}_i + b)y_i \geq 1, \ i = 1, ..., n \end{aligned} \tag{3.18}$$

where $\mathbf{w}$ is the normal vector to the hyperplane, $\mathbf{x}_i$ is the ith feature vector and y_i is its label. By solving the above quadratic programming problem, we obtain the separating hyperplane as $\mathbf{w}^T\mathbf{x}+b = 0$. In this work, the SVM implementation provided in MATLAB "Statistics and Machine Learning" toolbox is used. An attractive property of SVM is that we can use so-called kernel tricks to convert a non-linearly separable dataset to a linearly separable one. Some commonly used kernels are polynomial, Gaussian and radial basis function (RBF). After testing with different kernels, we found that the linear SVMs perform the best in cavernoma classification.

As mentioned in the imbalanced dataset handling section, multiple SVMs are trained in the proposed CAD system to detect cavernoma. As a consequence, a voting strategy is used in prediction. If the number of predictions of CM is larger than a threshold τ_v, the candidate block is deemed to be a CM. Otherwise, it is non-CM. The threshold τ_v is determined using the receiver operating characteristic (ROC) curve, as shown in Fig. 3.10. We can see the best value for voting threshold τ_v is 3.

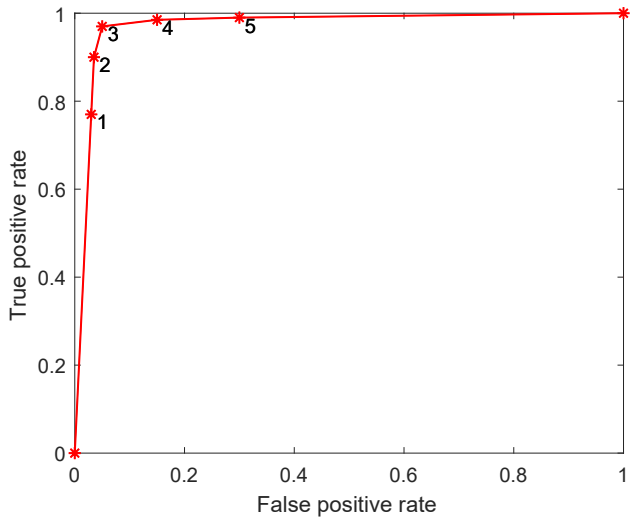


Figure 3.10: ROC for classification by SVMs.

3.4 Experimental Results

In this section, we evaluate the proposed CAD system for cavernous malformation detection based on two aspects: the accuracy of skull stripping, and the CM detection performance of the proposed system.

3.4.1 Efficiency of Skull Stripping

In this section, we evaluate the performance of the atlas-based skull stripping method in the proposed CAD system by calculating the dice coefficient of every slice. If B and G are the automatically extracted and manually extracted brain

images, respectively, the dice coefficient corresponding to (B, G) is obtained by using the following equation.

$$d(B, G) = \frac{2|B \cap G|}{|B| + |G|} \tag{3.19}$$

where $|B|$ and $|G|$ are the areas of B and G, respectively, and $|B \cap G|$ is the mutual area. Fig. 3.11 shows the distribution of dice coefficients for each patient and the average value is around 0.95. In other words, the atlas-based skull stripping method used in our CAD system can extract the brain region accurately, and no cavernoma region is excluded by the atlas-based skull stripping method. Fig. 3.12 (b) shows an example of the extracted brain. We also tried to use a morphology-based technique [54] in skull stripping. However, the morphology-based method does not perform very well and sometimes excludes cavernoma regions that are located at brain edge, as shown in Fig. 3.12 (c), while the atlas-based method can avoid the problem. The performance of the proposed skull stripping method is also compared with the BET [137]. Table. 3.1 shows the mean dice coefficients of the skull stripping results using our T2W image dataset. The proposed atlas-based method can outperform the BET in processing the T2W images with large slice thickness and a dice coefficient of 0.95 is obtained.

Table 3.1: Performance comparison of skull stripping methods.

Methods	BET [137]	Proposed
Dice Coefficient	0.90	0.95

3.4.2 System Performance

For subjective evaluation of the proposed CAD system, two examples of diagnostic results are shown in Fig. 3.13. It is observed that the atlas-based skull

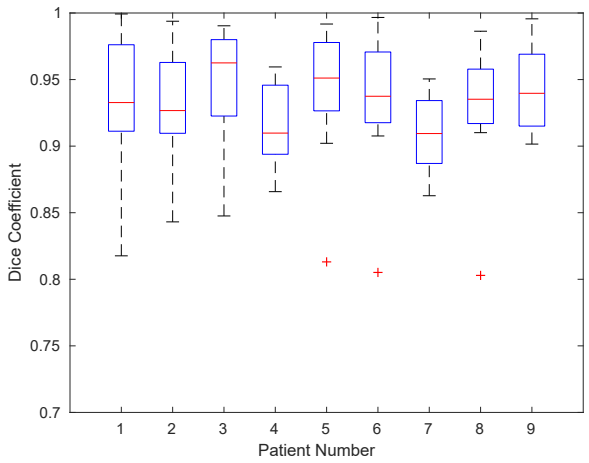
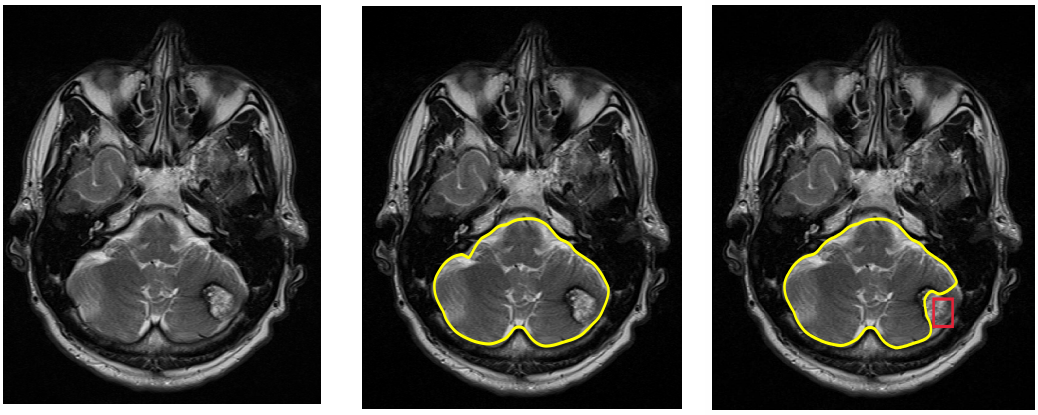


Figure 3.11: Boxplot of dice coefficients for all patients. Red lines denote the median dice coefficient and the '+'s denote the outliers.

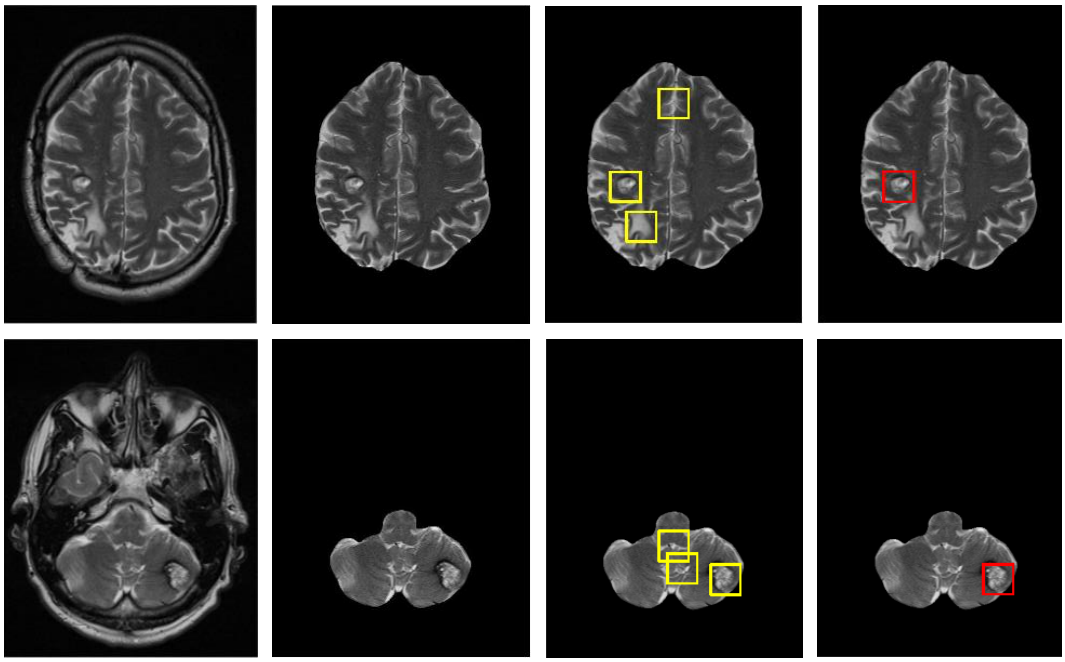


(a) (b) (c)

Figure 3.12: An example of skull stripping. (a) The input image, (b) extracted brain (in yellow boundary) using atlas-based method, (c) extracted brain (in yellow boundary) using morphology-based method. Note that in (c), the cavernoma region (shown in red box) is not included in the brain area.

stripping technique extracts the brain area precisely. The template matching technique produces three candidates in both examples. The SVM models correctly detect the cavernous malformations based on the features of each candidate region.

For a quantitative evaluation of the proposed SVM-based cavernomas detection system, we use the leave-one-out 'by patient' cross validation test [8], which means images of one patient is used for testing and the other patients



(a) (b) (c) (d)

Figure 3.13: Two qualitative results of the proposed system. (a) Two input images corresponding to the upper level (top row) and lower level (bottom row) brain slices, (b) extracted brain after skull stripping, (c) candidate cavernoma regions in yellow boxes after template matching, (d) detected cavernomas (in red boxes) after SVM classification.

Table 3.2: Parameters of training and testing.

Parameters	Value
No. of images	192
No. of candidate regions	1161
No. of CM regions*	207
No. of non-CM regions	954
No. of SVMs	5

*After oversampling by a factor of 9.

images are used for training the SVM models. After template matching, we obtain 1161 candidate regions consisting of 207 CM regions and 954 non-CM regions (see Table 3.2). For each patient, the test runs for 100 times and in each run, five SVMs (to offset the data imbalance) are learned based on the

randomly divided training dataset (so the prediction results may vary). After SVM classification, the sensitivity, specificity, and accuracy are calculated. Fig. 3.14 shows the average sensitivity, specificity, and accuracy for each patient. It is observed that, on average, 95% of CM regions could be identified and approximate 90% of false positives could be removed by the classification model.

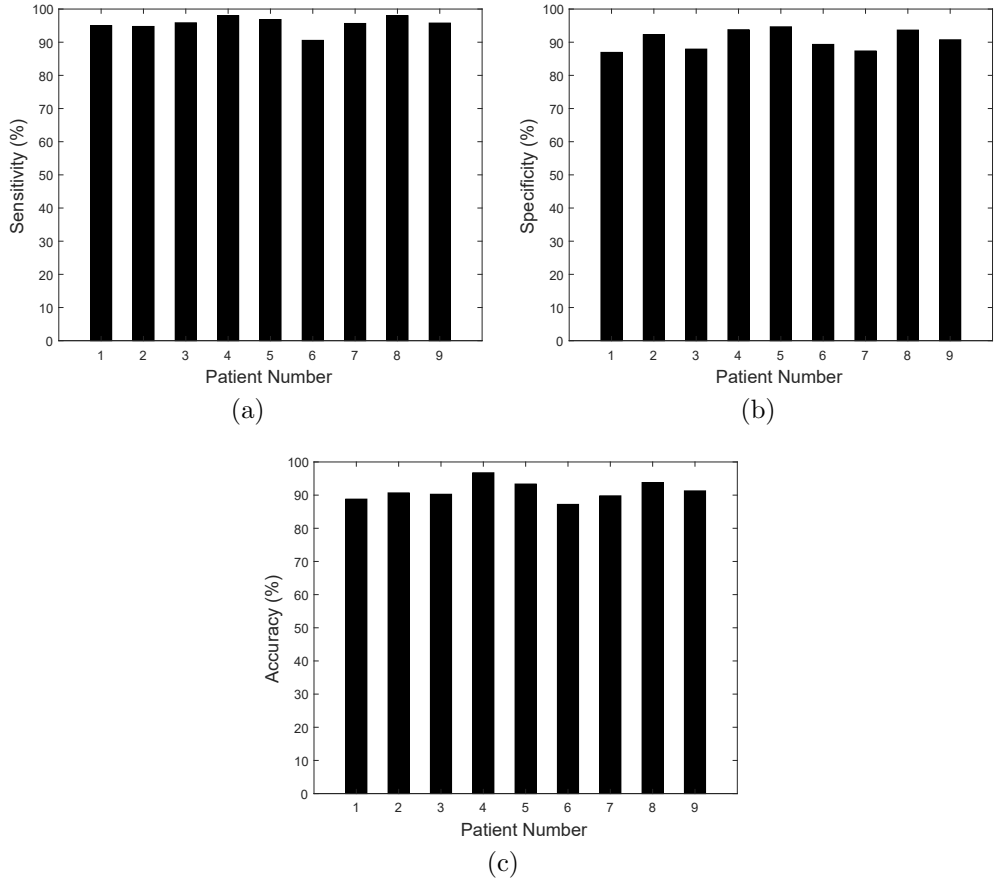


Figure 3.14: Performance of the CAD system for all patients. (a) Sensitivity, (b) specificity, and (c) accuracy

In the proposed technique, a feature vector $\boldsymbol{x}$ of length 965 is calculated for each candidate region, and the PCA is applied to reduce the feature dimensionality. To evaluate the significance of PCA (Section 3.3.3), the performance of the system is measured at different feature dimensionality, with results shown in Table. 3.3. It is observed that the performance (accuracy, sensitivity, and

specificity) is improved at a low dimensionality, which explains approximately 90% of the feature variance according to the cumulative variance plot in Fig. 3.15.

Table 3.3: Performance based on feature dimensionality

Dimensionality	Sensitivity(%)	Specificity(%)	Accuracy(%)
20	90	80	82
40	92	86	87
60	94	89	90
80	93	90	91
100	95	90	91
150	92	88	89
200	91	88	89
250	90	89	89
300	94	89	89
400	90	89	89
500	91	90	90
700	92	90	91

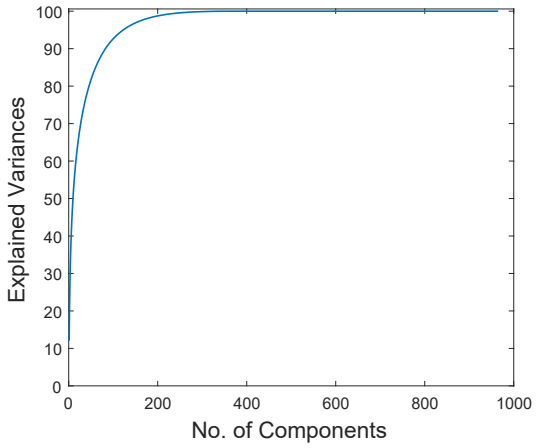


Figure 3.15: Cumulative variance plot.

Table 3.4 compares the performance of the proposed technique with three other existing techniques. The first two techniques [14, 40] were proposed for general tumor detection. To implement them for the cavernomas detection,

the candidate cavernoma regions are first obtained using the skull stripping and TM technique proposed in this paper. The classification of these candidate regions is then performed using the texture features and SVM classifier as in [14], and the DWT features with PCA and ANN classifiers as in [40]. The third technique [156], developed by us, is based on template matching and thresholding for the cavernomas detection. It is observed that the proposed technique outperforms the other three techniques with respect to all three performance criteria.

Table 3.4: Performance comparison.

Techniques	Sensitivity(%)	Specificity(%)	Accuracy(%)
Texture+SVM [14]	83	79	80
DWT+PCA+ANN [40]	88	82	82
TM+Thresholding [156]	86	80	81
Proposed	**95**	**90**	**91**

3.4.3 Computation Time

The proposed CAD system is implemented using MATLAB R2016b on a desktop computer with a 4.0 GHZ Intel i7-6700K CPU and 16.0 GB RAM. Typical processing time for one 3D image (with $512 \times 416 \times 20$ voxels) is approximately *2* minutes.

3.5 Conclusion

In this chapter, we have proposed a computer-aided diagnosis system for cavernomas based on axial T2W MRI images. The proposed system first extracts the brain region using an atlas-based skull stripping method. A template matching method is then used to detect the candidate cavernoma regions. Several textures, HOG and LBP features are extracted for each candidate,

and feature dimensionality is reduced using PCA. Finally, five SVM models are used to classify the candidate regions into cavernoma or non-cavernoma. Experimental results show promising performance for the proposed technique. The future study will be focused on the segmentation of cavernous malformations using high-resolution 3D MRI images and the evaluation of the system with a larger dataset.

Chapter 4

Automated Detection of Mesial Temporal Sclerosis

Mesial temporal sclerosis (MTS) is the commonest brain abnormalities in patients with intractable epilepsy. Its diagnosis is usually performed by neuroradiologists based on visual inspection of MRI scans, which is a subjective and time-consuming process with inter-observer variability. In order to expedite the identification of MTS, an automated computer-aided method based on brain MRI characteristics is proposed in this chapter. It includes brain segmentation and hippocampus extraction followed by calculating features of both hippocampus and its surrounding cerebrospinal fluid. After that, support vector machines are applied to the generated features to identify patients with MTS from those without MTS. The proposed technique is developed and evaluated on a data set comprising 15 normal controls, 18 left and 18 right MTS patients. Experimental results show that subjects are correctly classified using the proposed classifiers with an accuracy of 0.94 for both left and right MTS detection. Overall, the proposed method could identify MTS in brain MR images and shows promising performance, thus showing its potential clinical utility.

4.1 Introduction

Mesial temporal sclerosis (MTS) is the commonest brain structural abnormality in drug-resistant temporal lobe epilepsy patients [105], and is associated with gliosis and the loss of neurons in hippocampus [22]. In Fig. 4.1, an example of 3D reconstructed hippocampi from an anterior view of brain is shown. To evaluate the structural change of hippocampus noninvasively, MRI is the modality of choice to generate brain images, on which MTS characteristics of hippocampal atrophy and abnormal signals [135] are visually identified.

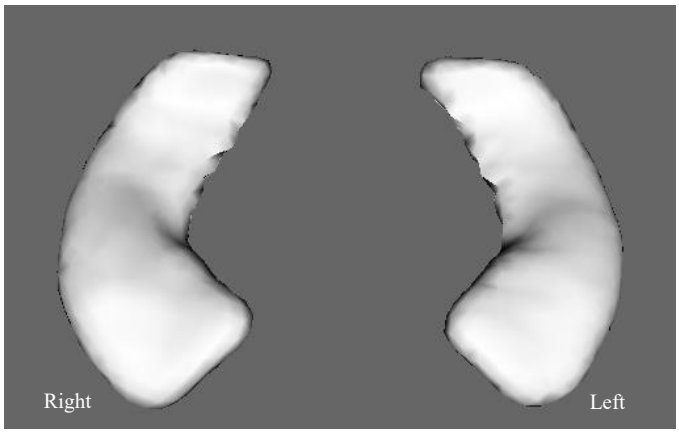
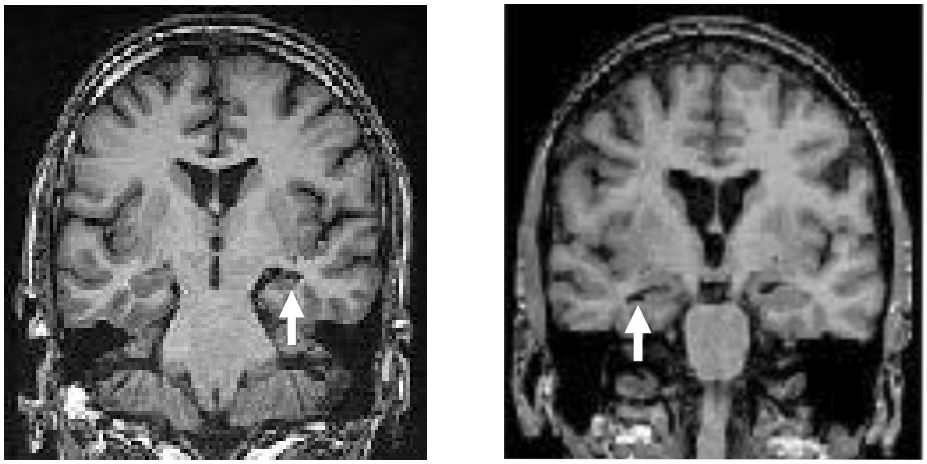


Figure 4.1: Left and right hippocampi in 3D.

The T1-weighted coronal MR images in Fig. 4.2 show the volume loss in the hippocampus with MTS, indicated by the white arrows. For cases (*e.g.* mild or bilateral MTS) that are difficult to be identified visually, the volumetric analysis can be helpful. This is typically performed manually on structural MR images with a high coronal resolution using the Cavalieri principle [51]. The hippocampal structures on each slice are outlined by a pre-clinical trainee and/or a research staff, and the areas from all slices are summed. Hippocampal volume is then obtained by multiplying the total area by the slice thickness. However, volumetric analysis is tedious and time-consuming, as it takes approximately 60 minutes to analyze one 3D image [59]. It requires expertise in brain anatomy and is therefore operator-dependent and subjective. For the

above reasons, volumetric analysis is restricted to tertiary care research facilities. It is reported that up to 60% of patients with MTS can be rendered seizure-free after temporal lobectomy [162], and this number increases if MTS is discovered at an early stage [98], when manual identification is challenging. It is therefore imperative to develop noninvasive techniques that can detect MTS efficiently and objectively.



(a) (b)

Figure 4.2: Two T1-weighted MRI coronal slices with hippocampal volume loss indicated by white arrows. (a) Left side MTS, (b) right side MTS. Note that the lost hippocampal volume is occupied by cerebrospinal fluid (CSF) and appears black in the indicated regions.

The hippocampal shape changes in patients with MTS have been studied by researchers and used in the detection of MTS. Hogan *et al*. [66] used the deformation-based hippocampal shape analysis to quantify the difference between the hippocampus with MTS and the normal hippocampus. They found the maximum shape changes in the medial and lateral hippocampal head and tail. Mumoli *et al*. [111] assessed the hippocampal shape changes in patients with MTS using shape analysis and surface reconstruction. The most prominent hippocampal atrophy was found in the CA1 subregion and subiculum. Kohan *et al*. [93] proposed to use hippocampal shape-based features to identify

patients with temporal lobe epilepsy.

Hippocampal volume loss is an important feature of MTS [46, 143], and has been used in identifying the patients with MTS. Bonilha *et al.* [19] proposed an automated MTS detection technique based on voxel-wise morphometry of the gray matter volume in hippocampus. A predefined threshold was used to identify the hippocampus with MTS. However, more advanced detection techniques usually combine multiple features within a machine learning framework. Jafari *et al.* [77] extracted the intensity, texture, and volume features of hippocampus from both T1-weighted and FLAIR images, and performed a linear classification to lateralize the mesial temporal lobe epilepsy. A similar study was proposed by Jafari [76] based on the quantitative analysis of the SPECT images. Focke *et al.* [47] used a linear SVM classifier with multiple features extracted from T1-weighted MRI, diffusion tensor imaging and T2 relaxometry images to detect MTS. Cantor *et al.* [23] proposed a technique for temporal lobe epilepsy detection using a linear SVM classifier in multi-parametric quantitative MR imaging, including T1 map, T2 map, fractional anisotropy and mean diffusivity. Rudie *et al.* [129] proposed to use machine learning and the morphological features of the whole brain for the detection of MTS.

Overall, some works have used shape features to analyze the difference in hippocampus between normal controls and patients with MTS. Some papers used multiple features, including the hippocampal volume, for the detection of MTS. However, the deformation of hippocampus may also be reflected by the volume change of surrounding cerebrospinal fluid [67], which, to the best of the authors' knowledge, has not been studied in the automated detection of MTS. In this chapter, we propose to use both hippocampus and CSF features with machine learning for MTS detection.

The remainder of this chapter is organized as follows. In Section 4.2, the materials of this study are introduced. In Section 4.3, the proposed methods

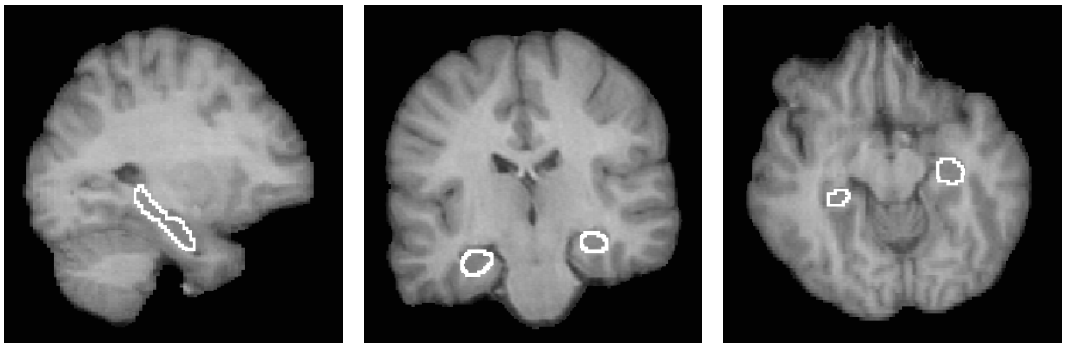
are presented. The experimental results are reported in Section 4.4. The chapter is concluded in Section 4.5.

4.2 Materials

In this study, the materials include 51 T1-weighted coronal MR images obtained from Henry Ford hospital [64, 69]. They are divided into 3 groups: (1) 18 images with left-side MTS, (2) 18 with right-side MTS, and (3) 15 images of normal controls. The three groups are respectively noted as LMTS, RMTS, and NC in this work. All of the images were initially reviewed by a neuroradiologist and then validated by a epileptologist epilepsy at the University of Alberta hospital. There was no disagreement between the two reviewers. These images were acquired using an inversion-prepared spoiled gradient recall sequence on either a 1.5T MRI scanner (TR/TI/TE = 7.6/1.7/500 ms, flip angle = 20°) with a scanning time of 5 minutes and 45 seconds or a 3.0 T (TR/TI/TE = 10.4/4.5/300 ms, flip angle = 15°) MRI scanner with a scanning time of 6 minutes. The voxel sizes are $0.78 \times 0.78 \times 2.0\ mm^3$ for 1.5T MR images, and $0.39 \times 0.39 \times 2.0\ mm^3$ for 3.0T MR images. In our dataset, 7 images in LMTS group, 10 images in RMTS group, and 5 images in NC group were obtained from the 3.0T MRI scanner. The other images (*i.e.* 11 LMTS, 8 RMTS, and 10 NC) were obtained on the 1.5T MRI scanner. Fig. 4.3 shows an image example with marked hippocampus in white boundaries on three view planes.

4.3 Methods

The overall schematic of the proposed MTS detection technique is shown in Fig. 4.4, and it consists of three main parts: (1) brain and hippocampus segmentation, (2) feature extraction, and (3) image classification. Details of



(a) (b) (c)

Figure 4.3: Example of an MR image with manually segmented hippocampus in three view planes. (a) Sagittal plane, (b) coronal plane, and (c) transverse plane.

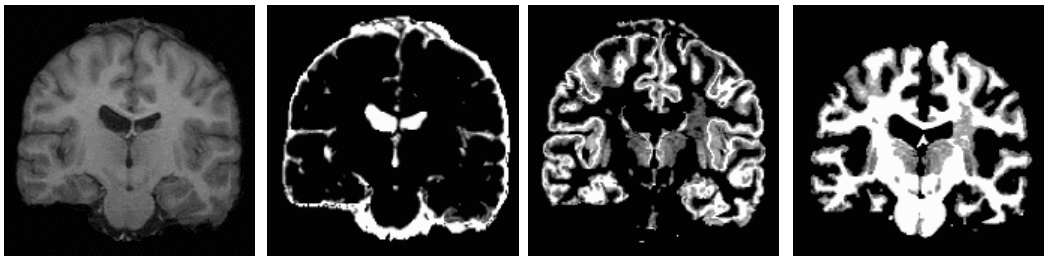
the three parts are given in Section 4.3.1 to Section 4.3.3.

Figure 4.4: The proposed MTS detection method.

4.3.1 Brain and Hippocampus Segmentation

As the shape change in hippocampus may affect the volume of the neighboring CSF, the brain MR image is first segmented into three tissue types (*i.e.* gray matter, white matter and CSF) using the "FAST" tool [169] from the FSL suite [138]. The FAST tool is based on a hidden Markov random field model and the expectation-maximization algorithm. The process is fast, reliable, and fully automatic. An example of the brain segmentation is shown in Fig. 4.5, and only the CSF partial volume estimation (PVE) is used in this work.

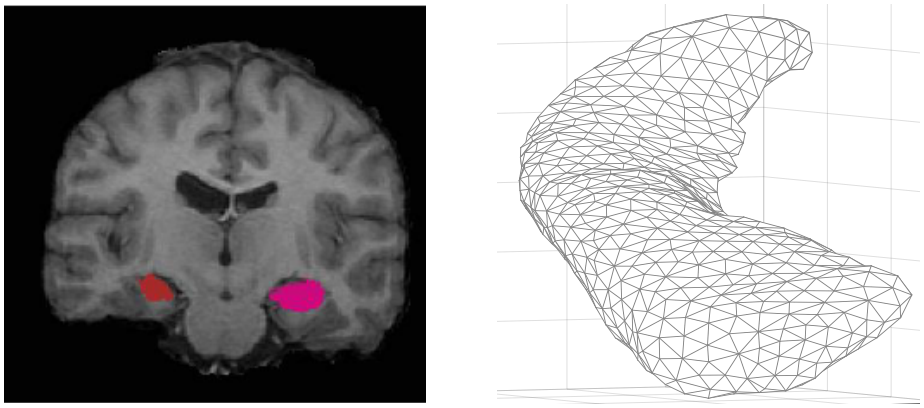
In this study, the extraction of hippocampus is performed using the FSL "FIRST" tool [120] that could automatically segment the brain sub-cortical structures based on a two-stage linear registration with standard brain templates and active appearance model within a Bayesian framework. The tool



(a) (b) (c) (d)

Figure 4.5: Segmentation results of one brain MR image. (a) Input image, (b) CSF, (c) gray matter, and (d) white matter.

generates the mesh of hippocampus and the transformation matrix from native to standard space. An example of hippocampus extraction is shown in Fig. 4.6.



(a) (b)

Figure 4.6: Hippocampus extraction results. (a) Extracted hippocampal area (the left and right hippocampus are in pink and red, respectively) in the coronal plane, (b) the mesh model for the right hippocampus.

4.3.2 Feature Extraction

Based on the results of brain and hippocampus segmentation, the hippocampus and CSF features are calculated as presented in the following sections.

Volumetric features

One feature of MTS is the hippocampal volume loss which is a common feature for MTS detection. Both the left hippocampal volume (LHV) and right hippocampal volume (RHV) are calculated from the segmentation results using the FIRST tool. They are then normalized to (*i.e.* divided by) the intracranial volume (ICV). The normalized hippocampal volumes are noted as $nLHV$ and $nRHV$ for the left- and right-side, respectively. The following two equations show the calculation of $nLHV$ and $nRHV$.

$$nLHV = \frac{LHV}{ICV} \times 100 \tag{4.1}$$

$$nRHV = \frac{RHV}{ICV} \times 100 \tag{4.2}$$

Typically, for an adult, the volume of each hippocampus is approximately 3.0 cm^3 compared with 1200 cm^3 for the intra-cranial volume [79, 141]. So, the $nLHV$ (or $nRHV$) value is approximately 0.25.

The volume asymmetry property of hippocampi is also useful in MTS detection, and calculated using the below equation.

$$HAI = \frac{2(LHV - RHV)}{LHV + RHV} \tag{4.3}$$

where HAI is the hippocampal volume asymmetry index.

Overall, the volumetric feature vector ($\mathbf{F_v}$) is defined as:

$$\mathbf{F_v} = \{nLHV, nRHV, HAI\}$$

Shape Features

In this study, the hippocampal shape is represented by the vertex coordinates of the hippocampus mesh, which is obtained in Section 4.3.1. All the coordinates are in a standard space (MNI152) so that the same location of hippocampi means the same vertex in the hippocampal mesh model. The shape

($\mathbf{S}$) of one hippocampus is represented as below.

$$\mathbf{S} = \{x^1, y^1, z^1, x^2, y^2, z^2, ..., x^n, y^n, z^n\} \tag{4.4}$$

where (x^i, y^i, z^i) are coordinates of the ith vertex and n is the number of vertices in each mesh. So, the shape of the hippocampus is represented by a vector of size $3n$.

Principal component analysis (PCA) is then applied to reduce the size of the shape vector. PCA transforms a set of possibly correlated variables into a set of uncorrelated low-dimensional variables (coefficients) using principal components. Dimensionality is reduced by only using the first few principal components, while retaining most of the variance in the dataset. The hippocampal shape ($\mathbf{S_i}$) of the ith image is represented using PCA as follows:

$$\mathbf{S_i} = \bar{\mathbf{S}} + \mathbf{p}\mathbf{F}_{\mathbf{s}i} \tag{4.5}$$

where $\bar{\boldsymbol{S}}$ is the mean hippocampal shape of the training images, $\mathbf{p}$ is the principal components, and $\mathbf{F}_{\mathbf{s}i}$ is the scores of principal components. $\mathbf{F}_{\mathbf{s}i}$ is used as a low-dimension feature vector of the ith image for MTS detection. In this study, the number of principal components is chosen such that at least 95% of the data variance is included in the principal components. After PCA, the shape features of left- and right-side hippocampi are noted as $\mathbf{F_{sl}}$ and $\mathbf{F_{sr}}$, respectively.

CSF-ratio Features

The CSF has been observed to encompass more space in a predefined box surrounding the hippocampus with MTS. To calculate the CSF-ratio features for each brain MR image, the following procedure is used.

a. The CSF PVE (*i.e.* obtained in Section 4.3.1) in native space is transformed to the standard MNI152 space using the transformation matrix from the FIRST tool.

b. Based on the standard brain atlas (*i.e.* MNI152), the bounding box of the left hippocampus is defined with the length, width, and height directions along the x, y, and z directions in the MNI space. The bounding box is then expanded outward by a certain distance (5 mm in our study) for each face. Its volume is noted as LBV.

c. Based on the defined box and CSF PVE in the standard space, the left-side CSF volume (LCV) is obtained. Therefore, the left-side CSF ratio (LCR) is calculated using $LCR = \frac{LCV}{LBV}$.

d. The right-side CSF volume (RCV) and the ratio (RCR) in a predefined box are calculated using the same procedure of [b - c] for the right hippocampus. Typically, both LCR and RCR are in the range of $[0.07, 0.11]$ for the surrounding box with 5 mm extensions. The extension is chosen at 5 mm in this study to maximize the correlation between the CSF-ratio features and the group labels.

e. The CSF volume asymmetry index is obtained as follows.

$$CAI = \frac{2(LCV - RCV)}{LCV + RCV} \tag{4.6}$$

The overall features of the CSF-ratio include LCR, RCR and CAI as follows.

$$\mathbf{F_c} = \{LCR, RCR, CAI\}$$

4.3.3 Classification

In classification, the support vector machine (SVM) is used for its efficiency and robustness [32, 123]. In this study, we use two SVM models - LSVM and RSVM for left and right MTS detection, respectively. Note that, the LSVM is trained using features $\mathbf{F} = \{\mathbf{F_v}, \mathbf{F_{sl}}, \mathbf{F_c}\}$ of normal controls and patients with left MTS. The RSVM is trained using features $\mathbf{F} = \{\mathbf{F_v}, \mathbf{F_{sr}}, \mathbf{F_c}\}$ of normal controls and patients with right MTS

During training, the SVM generates a hyperplane with the largest margin between the two classes. The objective function of SVM is shown below.

$$\min_{\omega,b} \ \frac{1}{2} \parallel \omega \parallel^2 + C \sum_{1}^{n} \xi_i, \tag{4.7}$$

where ω is a coefficients vector orthogonal to the hyperplane, ξ_i is the ith slack variable, and C is the regularization parameter. In this work, the SVM implementation provided in the MATLAB "Statistics and Machine Learning" toolbox is used.

The logistic regression (LR) and K-Nearest Neighbors (KNN) models are also evaluated for this classification problem. LR uses a linear combination of features in a logistic function to estimate the probability of the class label [84]. The linear coefficients are estimated using maximum likelihood estimation, and the classification results are obtained by thresholding the output probabilities. KNN uses the training data directly for classification instead of building a model [34]. It finds k nearest neighbors in training data based on the distance measure, and assigns the majority class label within this neighborhood to the testing data. For each classification model, the hyperparameters are optimized (using grid-search) in the training stage by minimizing a 10-fold cross-validation loss. Their classification performance are reported in the experimental results section.

4.3.4 Evaluation

To evaluate the MTS detection performance, the leave-one-out cross-validation strategy is used. One image is removed from training and used as the testing image. It is repeated until all images are tested. The LSVM model is evaluated using the images of normal controls and patients with left MTS, and the RSVM is evaluated using the images of normal controls and patients with right MTS. Sensitivity, specificity and accuracy are used as the performance measures for

MTS detection.

4.3.5 Performance Comparison

For comparison purposes, two existing MTS detection techniques are implemented and evaluated. The first technique [129] is implemented as follows. The brain is first segmented into 15 sub-cortical structures (*i.e.* left hippocampus, right hippocampus, left amygdala, right amygdala, left pallidum, right pallidum, *etc.*) using the FIRST tool. Note that [129] also used cortical regions, but these regions could not be obtained using the FIRST tool. The morphological features (*i.e.* volume, voxel intensity mean, and voxel intensity standard deviation) of these sub-cortical structures are then calculated and classified using SVM classifiers. The second technique [93] is implemented using linear SVMs based on hippocampal volume and shape features, including maximum shape diameter, shape volume, surface area, 3D compactness, *etc.* The comparison results are shown in the section of experimental results.

4.3.6 Feature Significance

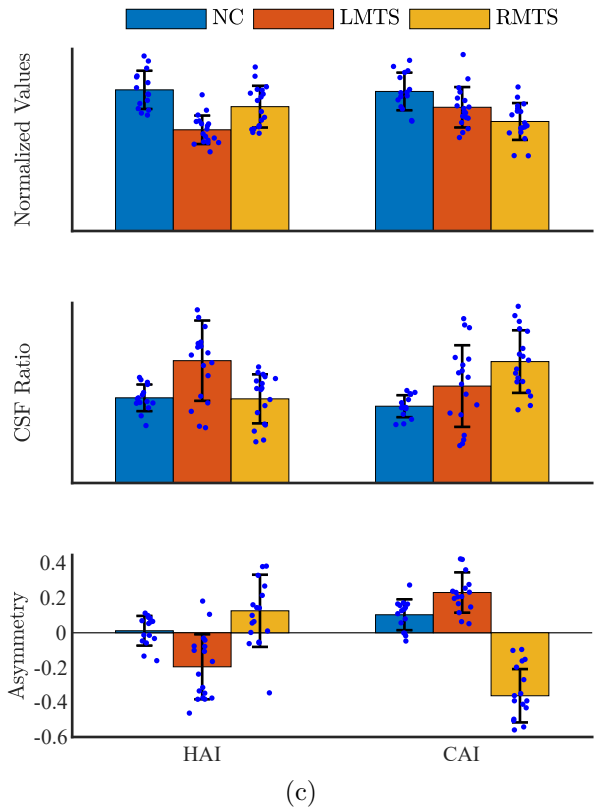
A statistical analysis for the hippocampal volume and the CSF-ratio features is performed to quantify the difference of each feature between control and patient groups, using a student's t-test [18] with a two-tailed significance level of 0.01. Fig. 4.7 shows the mean and standard deviation of each feature, and Table 4.1 shows the statistical analysis results (p values) between groups. It is observed that the LMTS group has the smallest mean value of $nLHV$, and there is a significant difference of $nLHV$ between LMTS and NC groups($p < 0.0001$), and between LMTS and RMTS groups ($p < 0.001$). For $nRHV$ feature, the RMTS group has the smallest mean value, and the difference between NC and RMTS groups is significant ($p < 0.0001$). Regarding the CSF ratios, as shown in Fig. 4.7(b), the LMTS group has the largest mean

value of LCR, and the RMTS group has the largest mean value of RCR. Significant differences are observed between NC and LMTS groups for LCR ($p = 0.02$), and between NC and RMTS groups for RCR ($p < 0.0001$). For the HAI feature in Fig. 4.7(c), the mean of the NC group is close to zero, while that of the LMTS group is negative and that of the RMTS group is positive. Note that the negative values on HAI indicate that $nLHV$ is smaller than $nRHV$, while positive values indicate that $nLHV$ is greater than $nRHV$. For CAI, there is no significant difference between NC and LMTS groups ($p = 0.464$), while the significant difference is observed in the other two group pairs(NC-RMTS and LMTS-RMTS).

Table 4.1: Statistical analysis of hippocampal volume and CSF-ratio features between groups (NC, LMTS, and RMTS). SD: standard deviation.

Features	p value		
	NC-LMTS	NC-RMTS	LMTS-RMTS
$nLHV$	**<0.0001**	0.123	**<0.001**
$nRHV$	0.128	**<0.0001**	0.034
HAI	**<0.001**	0.054	**<0.0001**
LCR	**0.002**	0.127	0.071
RCR	**0.002**	**<0.0001**	0.570
CAI	0.464	**<0.0001**	**<0.0001**

To visualize the shape deformations in hippocampi between normal controls and MTS groups, the mean hippocampal shape of the NC images group is first computed. Perpendicular displacements are then calculated between corresponding vertices from the mean shape to each subject. The t-statistic values [18] between the patient and control displacements are finally computed and visualized, as shown in Fig. 4.8. The mean left and right hippocampal shapes for the NC group are shown in Fig. 4.8 (a) and (c). The shape deformations of MTS hippocampi compared with normal controls are shown in Fig. 4.8 (b) and (d). Note that a positive t-statistic value means outward deforma-



(c)

Figure 4.7: Bar graph of mean volumetric and CSF-ratio features for three image groups (NC, LMTS, and RMTS). (a) Normalized hippocampal volume, (b) CSF-ratio, (c) hippocampal and CSF-ratio asymmetry index. Error bars are standard deviations. The scatter points (in blue color) show actual values of all subjects in each group.

tion, and a negative value means inward deformation. The individual color in Fig. 4.8 (b) shows the shape deformation in the right hippocampus between the RMTS and NC groups, and that in Fig. 4.8 (d) is the shape deformation in left hippocampus between the LMTS and NC groups. In both cases, obvious hippocampal sclerosis is noted on the lateral side. Similar findings were reported in the literature [9, 66].

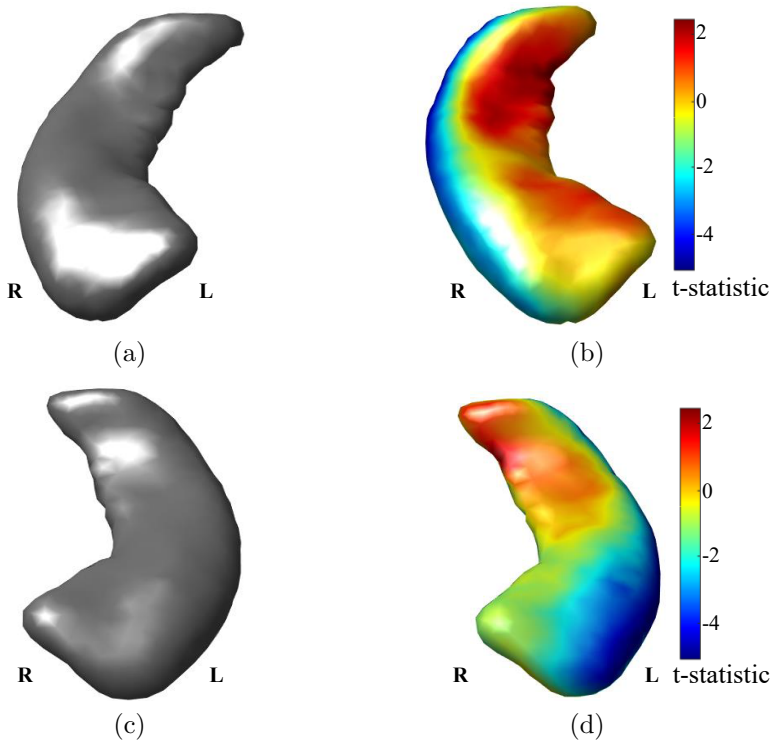


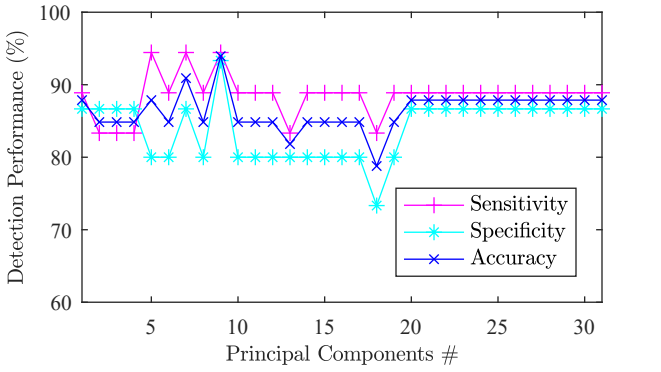
Figure 4.8: Shape deformations. (a) Mean right hippocampal shape of NC images, (b) the right hippocampal shape deformation for the RMTS group, (c) mean left hippocampal shape of NC images, (d) the left hippocampal shape deformation for the LMTS group. Note that the color stands for the t-statistic values.

4.4 Experimental Results

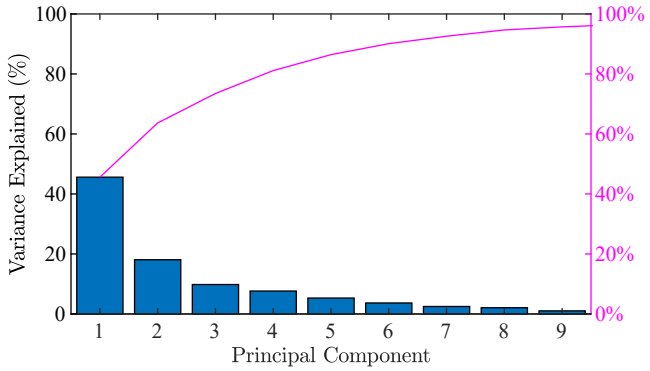
4.4.1 Effectiveness of PCA

In this work, the PCA is applied to the shape vector ($\mathbf{S}$) of the hippocampus. As we use the leave-one-out testing strategy, all images except one are used as the training set, and the omitted one is for testing. For PCA, the mean hippocampal shape and the principle components are calculated only using the training set. The relationship between the number of principal components used and the system performance is explored in this section. Fig. 4.9 (a) shows the performance of the left-side MTS detection using a varying number of principal components, while the scree plot in Fig. 4.9 (b) shows the percent variability explained by each principal component in decreasing order along

with the cumulated variance explained plot (the magenta line). It is observed that the first nine components could explain at least 95% of the total variance, and using more principal components does not help to improve the left MTS detection performance. Similar results are seen for right MTS detection. Overall, the number of principal components is chosen if they can explain at least 95% of the variance for both left and right side MTS detection.



(a)



(b)

Figure 4.9: PCA for left MTS detection. (a) Performance plot using different number of principal components, (b) scree plot of the percent variability explained by each principal component.

4.4.2 Classification Models

Table 4.2 shows the performance of MTS detection using different classification models: LR, kNN, and SVM. Since the leave-one-out testing strategy is used, multiple optimal hyperparameters are obtained for each classification model.

For kNN, the average number of neighbors k is 12. For SVM model, the majority is the linear SVM with an average regularization parameter C=32. The LR model has a ridge regularization with an average strength of 1 [65]. Overall, LR and SVM could achieve comparable performances, and better sensitivities than kNN for both left and right MTS detection. In this study, we choose to use SVM as the classification model in experiments.

Table 4.2: Performance comparison for different classifiers. Note: the values in parentheses are the standard errors estimated using binomial distributions.

	LMTS		
	Sensitivity	Specificity	Accuracy
kNN	0.83(0.09)	0.87(0.09)	0.85(0.06)
LR	0.94(0.06)	0.87(0.09)	0.91(0.05)
SVM	0.94(0.06)	0.93(0.07)	0.94(0.04)
	RMTS		
	Sensitivity	Specificity	Accuracy
kNN	0.83(0.09)	0.93(0.07)	0.88(0.06)
LR	0.89(0.08)	0.93(0.07)	0.91(0.05)
SVM	0.94(0.06)	0.93(0.07)	0.94(0.04)

In Fig. 4.10, the learned SVM weights are shown. The values in this figure are the mean of multiple SVM models obtained during the leave-one-out cross-validation process. Note that the features are normalized in training the SVM models. Fig. 4.10a shows the weights in LSVM. A large negative weight is observed for the left hippocampal volume ($nLHV$), which means the volume loss in the left hippocampus increases the classification value. The large positive weight for left-side CSF volume ratio (LCR) shows a positive correlation between CSF volume and MTS prediction. The first shape feature ($Shape1$) has a large weight at approximately -0.6, with the other shape weights are relative small. For Fig. 4.10b, the $nRHV$ has a negative weight at about -0.3, and the RCR has a positive weight at approximately +0.3. It illustrates the

decrease in $nRHV$ and the increase in RCR can increase the right MTS detection. The first shape feature is also observed to have a large contribution to the right MTS detection.

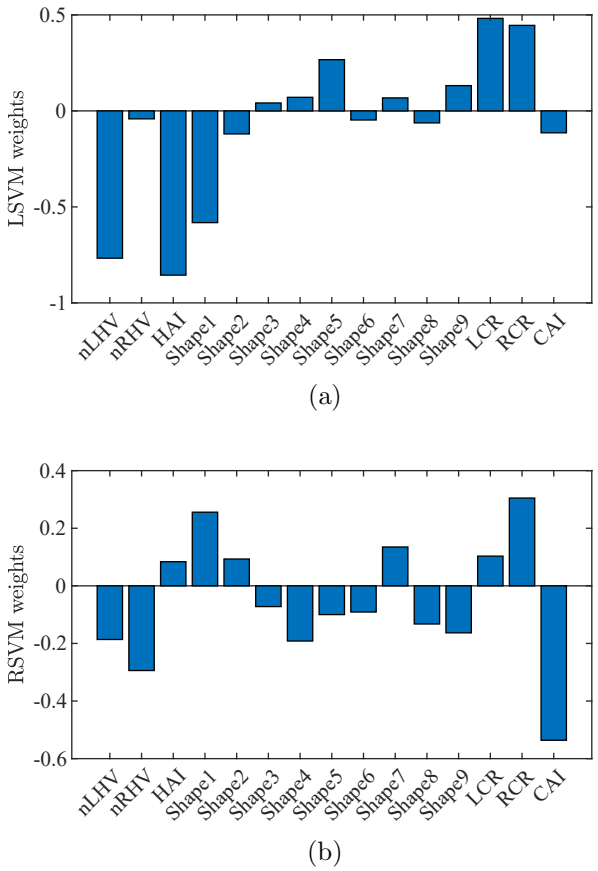


Figure 4.10: Bar plot of SVM weights. (a) LSVM, (b) RSVM.

4.4.3 System Performance

The confusion matrices for both LSVM and RSVM are shown in Table 4.3. For left MTS detection, 17/18 MTS images were detected, and 14/15 NON-MTS images were identified successfully, thus obtaining a sensitivity of 0.94, a specificity of 0.93, and an accuracy of 0.94. For right MTS detection, there was one false negative out of 18 RMTS images, and 14 healthy right hippocampi were identified correctly. Thus, the sensitivity, specificity, and accuracy of right MTS detection are 0.94, 0.93, and 0.94, respectively.

Table 4.3: Confusion matrix for LSVM and RSVM. Act.: Actual, Pred.: Prediction.

	LSVM		**RSVM**	
Act. \ Pred.	MTS	NONMTS	MTS	NONMTS
MTS	17	1	17	1
NONMTS	1	14	1	14

Fig. 4.11 shows the two miss detections (*i.e.* one in LMTS and another in RMTS) along with the segmented hippocampal areas. Fig. 4.11 (b) and (e) are the ground truth of hippocampal regions (in white contours), while (c) and (f) are the segmentation results (white regions) using the FIRST tool (in section 4.3.1). The segmentation of hippocampus with MTS is not accurate for these two cases, as the segmentation results include the lost hippocampal volume (indicated by the white arrows in (a) and (d)) as hippocampus regions. Improving the segmentation results of the hippocampus is our future research goal.

To plot the receiver operating characteristic (ROC) curves of SVMs, the distance to the hyperplane is converted to the posterior probability that an observation is classified into the MTS group using a sigmoid function [122]. By varying a predefined threshold of the probability, the true and false positive rates are calculated at each threshold value. Therefore, the ROC curves are plotted in Fig. 4.12 for both LSVM and RSVM. The area under the curve (AUC) measurement for LSVM is 0.97 and the AUC for RSVM is 0.98.

In Fig. 4.13, we compare the performance of each feature type (*i.e.* hippocampal shape, volume or CSF-ratio) in MTS detection. It is observed that the performance based on the CSF-ratio features is comparable (in left MTS detection) with or better (in right MTS detection) than that based on the hippocampal shape or volume features.

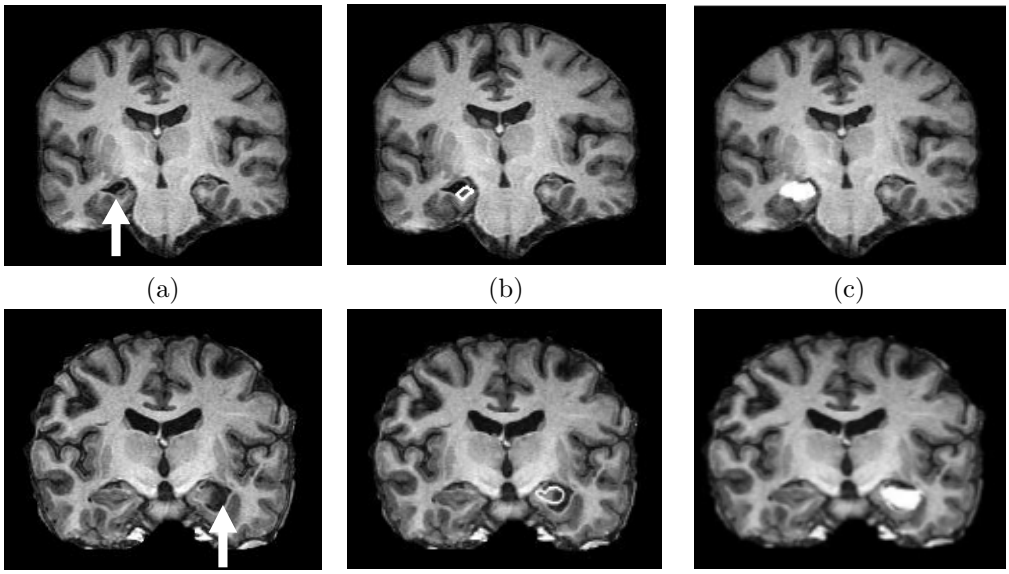


Figure 4.11: The two miss detections. (a) Right MTS indicated by the white arrow, (b) ground truth [64] of right hippocampus (in the white contour) for (a), (c) segmentation results of right hippocampus for (a), (d) left MTS pointed by the white arrow, (e) ground truth [64] of left hippocampus (in the white contour) for (d), (f) segmentation results of left hippocampus for (d).

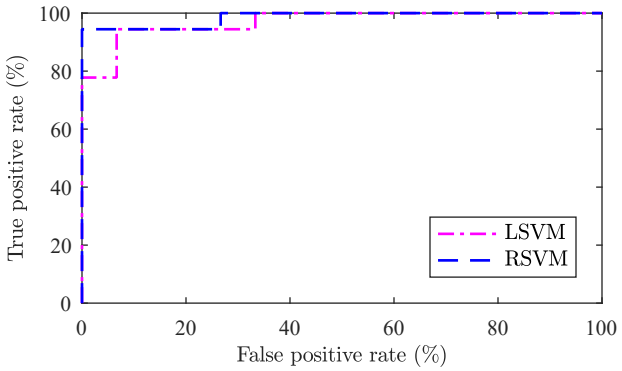
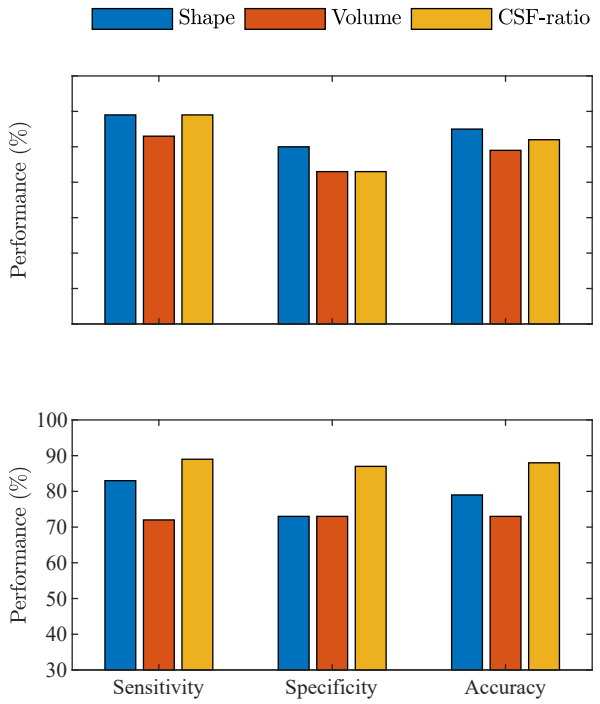


Figure 4.12: ROC curves for LSVM and RSVM

Table 4.4 shows the performance comparison results. It is observed that the proposed method could achieve a superior performance for MTS detection.



(b)

Figure 4.13: MTS detection using different features. (a) Left MTS detection, (b) right MTS detection.

Table 4.4: Performance comparison for different techniques. Sen.: sensitivity; Spe.: specificity; Acc.: accuracy.

	LMTS			**RMTS**		
Technique	Sen.	Spe.	Acc.	Sen.	Spe.	Acc.
Rudie [129]	0.89	0.87	0.88	0.83	0.87	0.85
Kohan [93]	0.78	0.80	0.79	0.78	0.87	0.82
Proposed	0.94	0.93	0.94	0.94	0.93	0.94

4.4.4 Computation Time

Regarding computation time, the proposed method takes approximately 7 minutes on average based on the time measurement on a computer with an Intel Core i5 CPU (2.6 GHZ) and 8 GB RAM, running MacOS. In this study, the FAST [169] and FIRST [120] tools are used for brain and hippocampus segmentation. Feature extraction and classification are implemented in MAT-

LAB R2017b. It is observed that the brain and hippocampus segmentation (*i.e.* preprocessing) takes over 95% of the overall computation time. Feature extraction and classification are computationally efficient as these together take approximately 10 seconds. Compared with manual hippocampal quantification, which takes approximately 1 hour for a typical 3D MR image [59], the proposed technique is much more time-efficient.

4.4.5 Discussion

Neuroimaging is the most commonly used tool in the identification of brain structural abnormalities. MTS characteristics in MR images mainly involve the reduced volume and abnormal morphometry of the hippocampus [75, 105]. This study aims to develop an automatic MTS detection technique using a machine learning approach based on the combination of the hippocampus and its surrounding tissue (*i.e.* CSF) features in T1-weighted structural MR images. Experimental results have shown a promising performance of the proposed technique, indicating its potential in clinical utility.

The three types of features (*i.e.* hippocampal volume, hippocampal shape and CSF features) used in this chapter could provide complementary information for MTS classification. Previous groups [42, 59, 129] have already shown that hippocampal volume and asymmetry properties are the most reliable features in identifying MTS or TLE. Some studies [66, 111] suggested that hippocampal shape analysis could provide further insights into MTS when there is little difference in the total hippocampal volume between normal controls and patients. Furthermore, the neuronal cell loss in the hippocampus may be reflected by the space change in cerebrospinal fluid surrounding the diseased tissue [67]. As seen in Fig. 4.13, CSF features are comparable with hippocampal volume and shape features in the performance of MTS detection.

Table 4.4 compares the proposed technique with two existing techniques

[93, 129]. Kohan *et al.* [93] used the hippocampal volume and shape features, and obtained a sensitivity of 0.78 for both left and right MTS detection. Rudie's technique [129] was implemented using the morphological features of 15 sub-cortical structures, and CSF properties were not included. It is observed that using a large number of structures did not achieve a better performance. One possible reason is that insignificant features were included and introduced noise to the classification problem. The results further illustrated the significance of CSF features in MTS detection.

The SVM is a classification model widely used in the detection of brain lesions, including tumors [58, 113], multiple sclerosis [102], stroke [33, 167] and arteriovenous malformations [50, 116]. One advantage of SVM is its robustness against the risk of over-fitting [72, 159]. Additionally, SVM could use kernel tricks to map features that are not linearly separable into a higher-dimensional space to make the separation much easier. However, this study finds that the linear SVM classifier outperforms the kernel-SVMs (polynomial and RBF), as reported in Table 4.2. This may be caused by the sample size, which may be not large enough to estimate the kernel-SVMs parameters. The classification problem in this study could also be considered as a multi-class classification problem using one classifier to classify an MR image into normal, left MTS, right MTS, or bilateral MTS. However, generating classification scores for the left and right hemispheres separately has been reported to be more clinically relevant [23, 129], which has been done in this study. Additionally, using two binary classifiers can detect bilateral MTS cases, even if the training dataset does not include bilateral cases.

This study can also be extended to the detection of some other neurological diseases, such as multiple sclerosis and Alzheimer's disease, which affect the hippocampal morphometry. Reduced hippocampal volume has been reported to have an association with the cognitive impairment in multiple sclerosis

[57] and Alzheimer's disease [140]. It is notable that neurological features of MTS are not limited to the change of hippocampal volume, shape, and CSF properties. MTS is also linked with whole-brain white matter and gray matter changes [46, 101]. Additionally, CSF biomarkers are also correlated with the neurodegenerative processes in MTS [143], multiple sclerosis [53, 57], and Alzheimer's disease [35]. Overall, combining both pathological and neuroimaging features may further improve the detection and prediction of neurological diseases.

Some limitations of our study are worth noting. First, the sample size is limited, which is a common problem in neuroimaging studies [23, 66, 111]. Therefore, leave-one-out cross-validation is used in this study to avoid the possibility of over-fitting and to estimate the general performance [27]. A larger dataset may contribute to developing even more robust and accurate classifiers for MTS. Second, this study only involves T1-weighted images, which are more sensitive to subtle abnormalities in brain tissues [20]. Features from T2 [23], FLAIR [77], PET [87], or SPECT [76] images have been reported to be helpful in MTS/TLE detection. Therefore, there is still room to improve the performance of the proposed system. In the future, a larger dataset including multiple modalities could be used to develop more comprehensive models.

4.5 Conclusion

In this chapter, a computer-aided detection technique for MTS is proposed based on both hippocampus and CSF features. Two SVMs are used for left- and right-side MTS detection, respectively. The experimental results demonstrate the promising detection performance of the proposed technique. In future, a larger dataset should be used to evaluate the proposed technique for patients with bilateral MTS, and to analyze the impact of MRI field strength on detection performance.

Chapter 5

Automated Detection of Focal Cortical Dysplasia

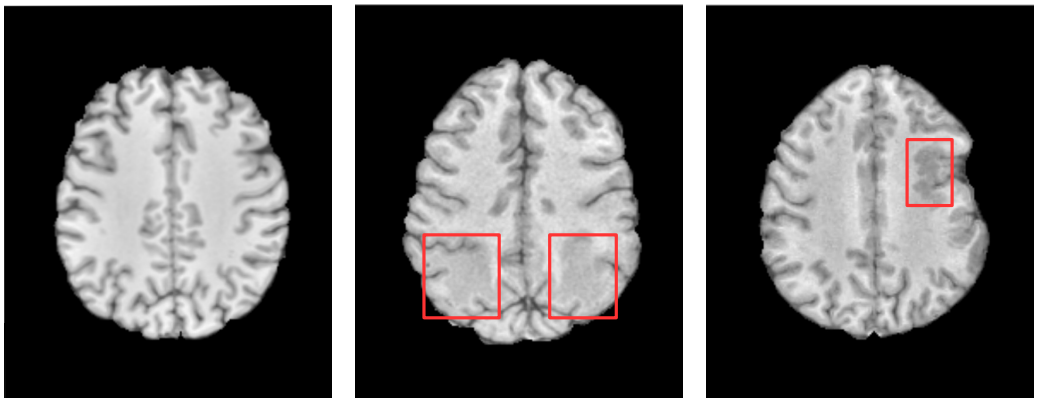
Focal cortical dysplasia (FCD) is one of the commonest epileptogenic lesions, and is related to malformations of the cortical development. The findings on MRI images are important for the diagnosis and surgical planning of FCD. In this chapter, an automated detection technique for FCD is proposed using MRI images and deep learning. The input MRI image is first preprocessed to correct the bias field, normalize intensities, align with a standard atlas, and strip the non-brain tissues. All cortical patches are then extracted on each axial slice, and these patches are classified into FCD and non-FCD using a deep convolutional neural network (CNN) with five convolutional layers, a max pooling layer, and two fully-connected layers. Finally, the false and missed classifications are corrected in the post-processing stage. The technique is evaluated using images of 10 patients with FCD and 20 controls. The proposed CNN shows a superior performance in classifying cortical image patches compared with multiple CNN architectures. For the system-level evaluation, nine of the ten FCD images are successfully detected, and 85% of the non-FCD images are correctly identified. Overall, this CNN based technique could learn optimal cortical (texture and symmetric) features automatically, and improve the accuracy of FCD detection.

5.1 Introduction

Focal cortical dysplasia (FCD) is the malformation of the cortical development, which may be caused by reasons of cortical architecture or cytological abnormalities [85]. It is the foremost cause of epilepsy in children and the third most significant cause in adults [100]. Magnetic resonance imaging (MRI) is widely used in identifying FCD as it can provide images of soft tissue with high contrast and resolution. In Fig. 5.1, three T1-weighted MR images (*i.e.* one healthy and two FCD images) are shown. The regions enclosed by red rectangular boxes in Figs. 5.1 (b) and (c) show two typical MRI features of FCD: blurred gray matter (GM) - white matter (WM) boundary and increased cortical thickness. Manual identification of FCD lesion on MR images is a time-consuming and subjective task even for an experienced specialist due to the subtlety of FCD lesions and complexity of brain anatomic structure [126]. In MR images, multiple tissue types (*e.g.* GM, WM, CSF) contribute to voxel/pixel intensities, which is known as partial volume effect, and it may also compromise the diagnosis of FCD. Therefore, it is important to develop a computerized system that can process brain MR images and detect subtle lesions automatically and objectively, to assist radiologists in analyzing images and making diagnosis.

Several techniques have been proposed for FCD detection in the literature. These techniques can be broadly divided into two categories: voxel-based morphometry (VBM), and surface-based classification (SBC). In the following, a brief review of techniques in these two categories is presented. Unless mentioned otherwise, T1-weighted MR images are assumed.

The VBM techniques [109] typically normalize images to a standard stereotactic space, segment the normalized images into regions of interest (ROIs), smooth these ROIs, and finally performs a voxel-wise statistical analysis to



(a) (b) (c)

Figure 5.1: T1-weighted MR images. (a) An axial slice of a healthy brain, (b) a FCD slice with blurred gray/white matter boundary, (c) a FCD slice with increased cortical thickness (typical values of cortical thickness for adults are between 1.5 mm and 3.0 mm [133]).

highlight significant anatomical differences between patients and healthy groups. The techniques usually generate statistical feature maps that are used to detect FCD regions by choosing an optimum threshold. Colliot *et al.* [29] proposed to use gray matter concentration (GMC) map for FCD detection. The GMC map is calculated by segmenting a normalized brain into GM, WM, and CSF, smoothing the GM mask, and calculating z-score [96] at each voxel in GM region. Note that z-score measures how many standard deviations a GMC value is above or below the mean of controls. A voxel is classified as FCD if its z-score value is greater than a predefined threshold. Similarly, Pail *et al.* [118] used GMC map to detect FCD within the temporal pole in patients with mesial temporal lobe epilepsy. Wagner *et al.* [148] proposed to use gray matter white matter junction (GWJ) and gray matter extension (GME) maps for FCD detection. The calculation of GWJ includes brain normalization, segmentation, binarization to obtain GM-WM junctions, smoothing and comparison with normal database. The GME map is calculated using the GM segment of brain. Wong-Kisiel *et al.* [163] further illustrated the effectiveness of GWJ and GME maps in FCD detection on a larger database of patients

and controls. House *et al.* [71] also used GWJ map, but based on T2-weighted MR images, to detect FCD regions.

The SBC techniques typically perform cortical reconstruction to obtain inner- and outer-cortical surfaces (using publicly available softwares such as FreeSurfer [49]), extract features at each vertex, and finally classify lesion vertices using the machine learning methods, such as artificial neural network (ANN) and support vector machine (SVM). Different surface-based features and classification models have been proposed in these SBC techniques. Besson *et al.* [17] calculated five features (*i.e.* cortical thickness, GM-WM blur, T1 hyper-intensity, sulcal depth, and curvature) at each vertex after cortical surface reconstruction, and performed both vertex-wise classification (using a bagged ANN) and cluster-wise classification (based on the statistic values of clusters). Hong *et al.* [68] used similar surface features, but vertex-wise classification was performed using Fisher linear discriminant analysis and cluster-wise classification was based on the statistical moments of clusters. Ahmed *et al.* [3] proposed to use bagged logistic regression to classify FCD vertices for "MRI-negative" patients. Adler *et al.* [2] proposed additional "doughnut" features (calculated on a circle region of vertices on the inflated surface) in the vertex classification using ANN. Clustering was applied in the end to remove false positive FCD clusters. Tan *et al.* [144] used surface features of both MRI and positron-emission tomography (PET) images, followed a two-step classification: a voxel-based SVM to maximize the sensitivity, and a patch-based classifier to remove the false positives. Jin *et al.* [82] proposed a FCD detection technique similar to [2], but with a larger database of images obtained from different epilepsy centers.

As observed, most techniques are based on either VBM or SBC. Typically, the VBM techniques use a few neurological features (*e.g.* GMC, GWJ, and GME), which are used for clinical diagnosis. Although, the VBM techniques

have shown a good performance, these techniques are sensitive to artifacts including misalignment, misclassification, and differences in anatomical structures [10]. On the other hand, the SBC techniques generate cortical surfaces, and calculate features that are robust to alignment and segmentation artifacts. These techniques are relatively efficient in FCD detection as they take into consideration the anatomical relationships across cortex [68]. But these techniques also have a high computational complexity because of the 3D surface reconstruction [127]. A further improvement may be possible by combining both voxel and surface features.

In recent years, deep learning techniques and especially convolutional neural networks (CNN) have shown great potential in image classification and segmentation problems since they could learn optimal features automatically [99]. Anthimopoulos *et al.* [7] proposed a lung pattern classification technique for interstitial lung disease based on a deep CNN architecture of 5 convolutional layers with 2×2 kernels, leaky rectifier linear unit (LReLU), and average pooling. The CNN architecture is trained and evaluated on non-overlapping image patches, and it shows superior classification performance over feature-based techniques (e.g., intensity, texture, LBP features). Pereira *et al.* [121] proposed a brain tumor segmentation technique based on a CNN architecture with 4-6 convolutional layers, 3×3 kernels, and LReLU. To the best of the authors knowledge, deep learning has not been applied to the detection of FCD. In this chapter, we propose an automated FCD detection technique using a deep CNN architecture. The rest of this chapter is organized as follows. In Section 5.2, the materials used in this work are introduced. The proposed methods are presented in Section 5.3, and the experimental results are reported in Section 5.4. Section 5.5 concludes this chapter.

5.2 Materials

In this study, the patient group includes a retrospective cohort of 10 patients who underwent T1-weighted imaging on the 1.5T Siemens MRI scanner at the University of Alberta Hospital, with confirmed FCD lesions. The study was approved by the Alberta health services and University of Alberta research ethics board. The MR images were acquired using the T1-weighted magnetization-prepared rapid-acquisition gradient echo (MPRAGE) sequence (TR = 2130ms, TE = 3.91ms, and flip angle = 15°) with an isotropic voxel dimension of 1×1×1 mm^3. Each 3D MR image has 208×256×192 (coronal×sagittal×axial) voxels. The FCD lesions are labeled manually on axial slices of the 3D MR images by an expert neurologist at the University of Alberta Hospital.

The control group includes 10 healthy subjects and 10 temporal lobe epilepsy (TLE) patients. The healthy controls were obtained from the IXI dataset [25], and the T1-weighted images were acquired on a 3.0T Philips MRI scanner at Hammersmith Hospital (TR = 9.6ms, TE = 4.6ms, and flip angle = 8°) with a voxel dimension of 0.9×0.9×1.2 mm^3 and an image resolution of 256×256×150. The TLE controls were obtained from the Henry Ford Hospital (HFH) dataset [64, 69], and the T1-weighted images were acquired on a 1.5T GE MRI scanner (TR = 7.6ms, TE = 500ms, and flip angle = 20°) with a voxel dimension of 0.8×0.8×2 mm^3 and an image resolution of 256×256×124. All of the control images were reviewed by an neurologist at the University of Alberta, and confirmed with no FCD lesions.

5.3 Methods

The overall schematic of the proposed technique is shown in Fig. 5.2. It includes four modules: 1) preprocessing, 2) patch extraction, 3) deep learning classification, and 4) post-processing. Details of the proposed methods are

presented in the following sections.

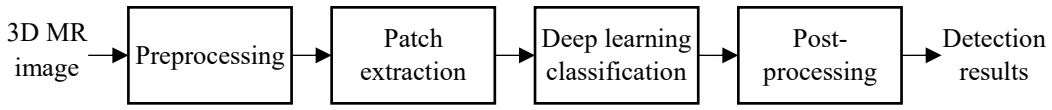


Figure 5.2: Schematic of the proposed FCD detection method.

5.3.1 Preprocessing

T1-weighted MR images are preprocessed using FMRIB Software Library (FSL) v5.0 [81], which involves the use of several FSL tools. First, the bias field correction algorithm in the automated segmentation tool [169] is applied to correct intensity non-uniformities caused by the low-frequency artifacts. Second, the linear image registration tool [80] is performed to align the image with a standard brain atlas [48]. The brain extraction tool [137] is then used to remove the non-brain tissues from the whole-head image. Finally, intensity normalization is applied to keep the image intensity in the range [0, 1].

5.3.2 Patch Extraction

Since FCD is a malformation of cortical development, neurologists are typically interested in the cortical regions and GM-WM junctions. Therefore, we extract rectangular patches of brain cortex on axial MR slices for FCD detection. The patch extraction is performed by first extracting patch masks on the gray matter probability map (GMPM) of the brain atlas [48], and then applying the masks to the preprocessed brain images. Patch masks are extracted using the following procedure.

1) Each axial slice of GMPM is divided into horizontal strips of height h ($h = 16$ pixels in this work). An example of GMPM division is shown in Fig. 5.3(a), and a horizontal strip is shown in Fig. 5.3(b).

2) On each horizontal strip, the Radon transform [36] is performed vertically to obtain the projection onto the horizontal axis. Fig. 5.3(c) shows the line plot of the Radon transform of the horizontal strip in Fig. 5.3(b).

3) To locate the cortex region in a horizontal strip, the values of its Radon transform are checked from left to right. Once the value changes from 0 to a positive value, a rectangular box of size $h \times w$ ($w = 32$ pixels) is placed at this point to be one patch mask. This is shown as the left green box in Fig. 5.3(d). Based on the symmetric property of the brain atlas and the mirror reflection about the middle sagittal line, the right-side patch mask is obtained, shown as the right green box in Fig. 5.3(d).

4) Cortical patches along the inter-hemispheric fissure are extracted, shown as the blue rectangles in Fig. 5.3(e).

5) Repeat step 1) - 4) on every axial slice of GMPM, the patch masks of the whole brain are extracted.

For a preprocessed image, the cortical patches are extracted based on the patch masks obtained from the above procedure. One example slice is shown in Fig. 5.4, and the extracted image patches are shown as the green and blue rectangular regions. In order to capture both the texture and symmetric properties of cortex, we use a pair of patches as the input to a classification model. Each cortical patch is paired with the mirror reflection of the corresponding patch on the other hemisphere. For example, in Fig. 5.4, patch pair 1 is formed by the image patch P_1 and the reflection of patch P_2 (noted as $\bar{P}_2$), and patch pair 2 is formed by P_2 and the reflection of P_1 (noted as $\bar{P}_1$). The patch pair has the same label as the image patch. Therefore, the axial slice in Fig. 5.4 could generate 42 patch pairs since there are 42 patches in this slice.

Using the above procedure, 129,780 (30 images $\times$ 4326 patches/image)

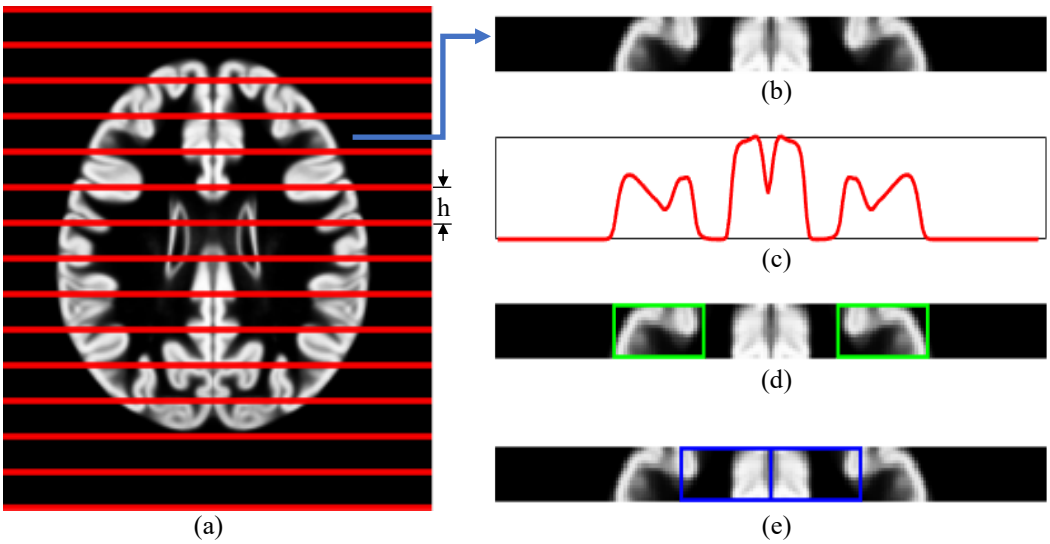


Figure 5.3: Patch masks extraction on one axial slice of GMPM. (a) The axial slice divided into horizontal strips of height h, (b) one horizontal strip, (c) line plot of the Radon transform of (b), (d) two external patches (the green boxes), and (e) two middle patches (the blue boxes) along the inter-hemisphere fissure. Note that all patch masks have the same size of $h \times w$.

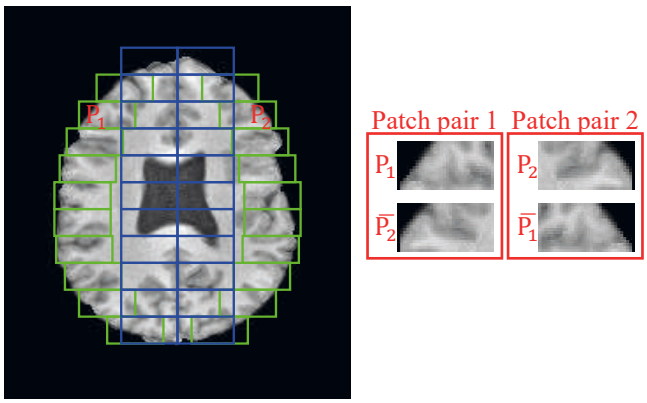


Figure 5.4: Patches extracted on one axial slice.

patches are extracted from the dataset (of 30 MR images) in this study. Out of these, 2,302 patches are labeled as FCD and 127,478 as non-FCD. To overcome the imbalance between FCD and non-FCD patches, data augmentation is applied. The traditional affine transform based data augmentation is not suitable for the cortical dysplasia images because the texture properties of a lesion region may be changed inappropriately after transformation. In this study, we

use up-sampling method to increase the number of FCD patches, which is achieved by re-extracting FCD patches starting from the ith ($2 \leq i \leq h$) row on each axial slice. This results in a total of 36,786 FCD patches (with an up-sampling factor of approximately 16). The same number of non-FCD patches are selected randomly to balance the two classes. All patches are divided into training, validation, and testing sets at ratios of 70%, 15%, and 15%, respectively.

5.3.3 Deep Learning Classification

Deep learning techniques, especially CNNs, have been used extensively in image classification problems due to their superior performance. In this chapter, a CNN architecture is proposed for cortical patch classification, as shown in Fig. 5.5. The input is one patch pair, and output is the classification result (*i.e.* probability of being FCD). The CNN architecture consists of five convolutional layers (Conv1 - Conv5), one pooling layer, and two fully-connected layers (FC1 and FC2). The details of this architecture are listed in Table 5.1, and the following types of CNN layers are involved.

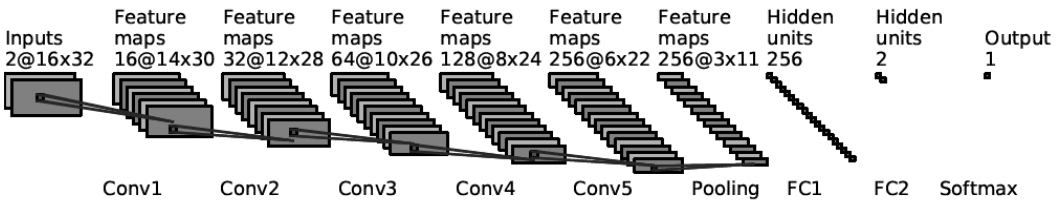


Figure 5.5: Schematic of the proposed CNN architecture (with 5 convolutional layers).

1) *Convolutional layers:* In a convolutional layer, a neuron is only connected to a local area of input neurons instead of full-connection so that the number of parameters to be learned is reduced significantly and a network can grow deeper with fewer parameters. In the proposed CNN,

Table 5.1: Details of the proposed CNN architecture with 5 convolutional layers, 3×3 filters, and scale parameter $k = 4$.

Layer	Type	Filters (#, size)	Activations	Learnables
Inputs	Input	-	16×32×2	
Conv1	Convolution	16, 3×3	14×30×16	Weight: 3×3×2×16 Bias: 16
	Normalization	-	14×30×16	Offset: 16 Scale: 16
	ReLU	-	14×30×16	-
Conv2	Convolution	32, 3×3	12×28×32	Weight: 3×3×16×32 Bias: 32
	Normalization	-	12×28×32	Offset: 32 Scale: 32
	ReLU	-	12×28×32	-
Conv3	Convolution	64, 3×3	10×26×64	Weight: 3×3×32×64 Bias: 64
	Normalization	-	10×26×64	Offset: 64 Scale: 64
	ReLU	-	10×26×64	-
Conv4	Convolution	128, 3×3	8×24×128	Weight: 3×3×64×128 Bias: 128
	Normalization	-	8×24×128	Offset: 128 Scale: 128
	ReLU	-	8×24×128	-
Conv5	Convolution	256, 3×3	6×22×256	Weight: 3×3×128×256 Bias: 256
	Normalization	-	6×22×256	Offset: 256 Scale: 256
	ReLU	-	6×22×256	-
Pooling	Max Pooling	-	3×11×256	-
FC1	Dropout	-	8448	-
	Fully-connected	-	256	Weight: 8448×256 Bias: 256
FC2	Fully-connected	-	2	Weight: 256×2 Bias: 2
Outputs	Softmax	-	1	

each convolutional layer consists of three operations: convolution, batch normalization, and ReLU activation.

(a) *Convolution* The output of a convolution operation is computed

by convolving the input with a number of filters, as shown in the following equation:

$$\mathbf{x}_j = \mathbf{I} * \mathbf{W}_j + b_j, j = 1, 2, ..., N \tag{5.1}$$

where $\mathbf{I}$ is the input, N is the number of filters, $\mathbf{x}_j$ is the output corresponding to the jth convolution filter, $\mathbf{W}_j$ is the weights of the jth filter, and b_j is the jth bias. In the first convolutional layer of the network in Fig. 5.5, $\mathbf{I}$, $\mathbf{W}_j$, and $\mathbf{x}_j$ have dimensions of $16\times32\times2$, $3\times3\times2$, and 14×30, respectively. Note that $\mathbf{x}_j$ does not include any partial convolution output. As a result, the feature map dimension decreases after each convolution operation. In CNN, the number of filters and the filter size are two hyperparameters that need to be specified. The number of filters, N, at each convolutional layer, is defined as follows:

$$N_l = k * 2^{l+1}, l = 1, 2, ..., L \tag{5.2}$$

where N_l is the number of filters at the lth convolutional layer, L is the total number of convolutional layers (*e.g.* $L = 5$), and k is a scale parameter.

(b) *Batch normalization:* After convolution operation, batch normalization is used to speed up the training of CNN and reduce the effect of initialization [74]. It normalizes each input channel across a mini-batch by subtracting mean of the mini-batch and dividing by its standard deviation. The normalized values are then scaled and shifted as follows:

$$y_i = \sigma\hat{x}_i + \beta \tag{5.3}$$

where y_i is output value, $\hat{x}_i$ is the normalized input value, σ and β are the scale and offset factors that are learnable during the network training.

(c) *ReLU:* Nonlinear activation functions are applied element-wise to increase the non-linear properties of the network. In the CNN, rectified linear unit (ReLU) is the most commonly used nonlinear activation function [112], and it replaces all negative pixel values by zeros using the following equation:

$$f = \max(0, y) \tag{5.4}$$

where y is the input value to ReLU, and f is the output. Compared with *tanh* and *sigmoid* activations, ReLU has the advantages of speeding up training [97] and inducing the sparsity in hidden units [56].

Step (a)-(c) are repeated five times (as shown in Table 5.1). After the ReLU module, the feature map f is used as the input (I) of the next convolutional layer (if appropriate). After the five convolutional layers in the proposed architecture, 256 features maps, each of dimension 6×22, are obtained, as shown in Table 5.1. These are then down-sampled in a pooling layer.

2) *Pooling:* In a pooling layer, the feature maps are down-sampled using an average or a max operation to reduce the number of parameters and control overfitting in a network. In this study, a max pooling layer with filter of size 2×2 with a stride of 2 is used. Before the pooling, there are 256 feature maps of size 3×11. After the pooling, the feature maps are flattened and concatenated to form a 8448×1 feature vector.

3) *Fully-connected (FC) layers:* The proposed CNN has two FC layers as shown in Fig. 5.6. After pooling and flattening, the feature vector of size 8448×1 is fully connected to 256 nodes at the first FC layer. The second

FC layer has 2 nodes, and its output scores (noted as $z_i,\ i = 1, 2$) are fed to a softmax layer where the classification probability is calculated.

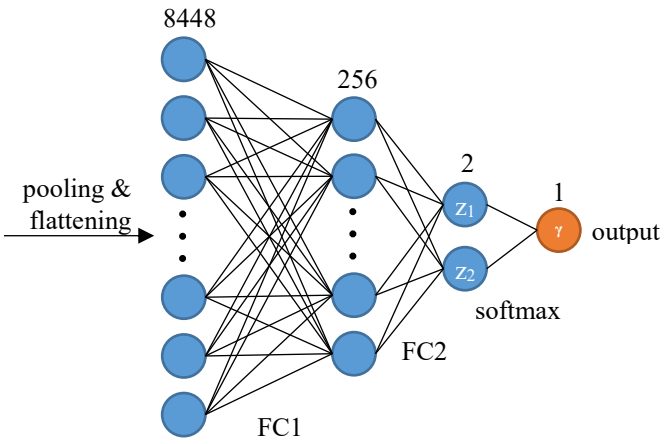


Figure 5.6: Fully-connected layers.

As observed in Table 5.1, the proposed network has a total of 2,556,914 parameters (numbers of parameters listed in the last column) that need to be updated during the training process, and the first FC layer takes approximately 85% of this number. To speed up training [139], dropout is used in the first FC layer removing some nodes with a probability ($p = 0.4$) during training. In other words, 40% of the 8448×256 weights (chosen randomly) are not updated at each training iteration.

4) *Softmax:* The softmax function is applied at the last layer to generate the probability distribution of the classification results. The predicted probability is computed using the following equation.

$$\gamma = \frac{e^{z_1}}{e^{z_1} + e^{z_2}} \tag{5.5}$$

where γ is the probability that the input belongs to the FCD class.

Training: To train the CNN, the loss function and the optimization algorithm must be defined. In this study, we use the cross entropy loss function

for multi-class classification, which is defined as follows:

$$E = -\sum_{s=1}^{S} t_s \ln \gamma_s \tag{5.6}$$

where S is the number of training samples (*i.e.* image patch pairs), t_s is the ground truth value that the sth sample belongs to the FCD class, and γ_s is the output probability of the network. To minimize the loss function E, the stochastic gradient descent with momentum (SGDM) [128] optimizer is used in the proposed technique. The SGDM accelerates stochastic gradient descent (SGD) and reduces the oscillation problem of SGD by adding the contribution from the previous iteration to the current iteration. The update to the network parameters is performed using the following equation:

$$\theta_{n+1} = \theta_n - \alpha \bigtriangledown E(\theta_n) + \mu(\theta_n - \theta_{n-1}) \tag{5.7}$$

where n is the iteration number, θ is the parameter vector, $\bigtriangledown E(\theta)$ is the gradient of the loss function, α is the learning rate, and μ is momentum. In this work, the momentum is set to 0.9. The learning rate is initialized to 0.01, and drops every 10 epochs by a factor of 0.2. The mini-batch size is set to 128 and the maximum number of epochs is set to 40.

5.3.4 Post-processing

A patch is considered as FCD if $\gamma > \tau$, where τ is a threshold that maximizes the validation accuracy. In the post-processing stage, the classification results of both a patch and its four neighbors are considered. As shown in Fig. 5.7, the four neighbors of a patch P include two patches (N_1, N_2) from the same slice, one patch (N_3) from the upper slice, and one (N_4) from the lower slice. The following two steps are performed to remove noise and avoid misclassifications. 1) If P is classified as FCD and N_1-N_4 are classified as non-FCD, P is treated as a noise and re-labeled as non-FCD. 2) If P is a non-FCD patch and N_1-N_4

are FCD patches, P is treated as a misclassification and re-labeled as FCD. After post-processing, an image with FCD is corrected identified if the detected FCD patches and the ground-truth FCD regions have overlap.

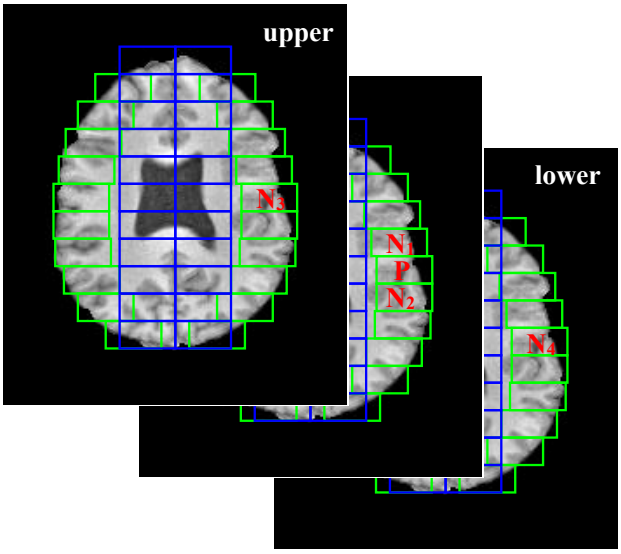


Figure 5.7: A patch (P) and its four neighbors (N_1-N_4).

5.4 Experimental Results

In this section, we first present the implementation of the proposed technique. The evaluation results of the proposed CNN model are then provided. Finally, the system-level evaluation is presented.

5.4.1 Implementation

The proposed technique is implemented using MATLAB R2018a with the neural network toolbox. All experiments are performed on a Windows 10 computer with Intel i7-6700K CPU and 16 GB RAM. The CNN architectures are trained using a NVIDIA GeForce GTX 970 graphic card.

5.4.2 Evaluation of the CNN

In this section, the classification performances of multiple CNN architectures are evaluated and compared with the proposed network. These architectures are constructed based on variations of the hyperparameters (*i.e.* number of layers, number of filters, type of activation function, and dropout probability). Each architecture is trained on the training dataset, and fine tuned on the validation dataset. The performance of each architecture is evaluated on the testing patch dataset using accuracy, which is calculated using the following equation.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{5.8}$$

where TP, FP, TN and FN are the numbers of true-positive, false-positive, true-negative, and false-negative patches, respectively.

Table 5.2 compares multiple CNN architectures in the patch classification performance. In this table, the last row in bold shows the testing accuracy (0.941) of the proposed architecture. The CNNs using 4 and 6 convolutional layers obtain a classification accuracy of 0.892 and 0.905, respectively, which are roughly a 5% drop in accuracy. Using smaller (2×2) filters reduces the performance by about 5%, and using larger (4×4) filters by another 1%. It illustrates that the optimal size of the receptive field corresponds to the convolutional filters of size 3×3. To determine the parameter k, values of 3 and 5 are tested, and it is observed that neither decreasing nor increasing the number of filters could increase the performance. LReLU activations with scale factors 0.01 and 0.1 are also evaluated with the proposed CNN, and they are inferior to the use of the ReLU activation function. The use of a dropout layer with a probability of 0.4 could increase the classification accuracy from 0.897 to 0.941 comparing with the CNN without dropout. Using either a lower (0.3) or a higher (0.5) dropout probability results in a performance drop.

Table 5.2: Performance of different CNN architectures. Acc.: accuracy.

# layers	Filter size	# filters (k)	Activation function	Dropout (p)	Testing acc.
4	3×3	4	ReLU	0.4	0.892
6	3×3	4	ReLU	0.4	0.905
5	2×2	4	ReLU	0.4	0.897
5	4×4	4	ReLU	0.4	0.881
5	3×3	3	ReLU	0.4	0.903
5	3×3	5	ReLU	0.4	0.914
5	3×3	4	LReLU(0.01)	0.4	0.930
5	3×3	4	LReLU(0.1)	0.4	0.918
5	3×3	4	ReLU	0	0.897
5	3×3	4	ReLU	0.3	0.911
5	3×3	4	ReLU	0.5	0.930
5	**3×3**	**4**	**ReLU**	**0.4**	**0.941**

Fig. 5.8 shows the receiver operating characteristic (ROC) curve for the proposed CNN. It illustrates the true positive rate against false positive rate based on the classification results on the testing dataset. The area under the ROC curve is 0.985, which indicates a good classification performance [106] of the proposed CNN.

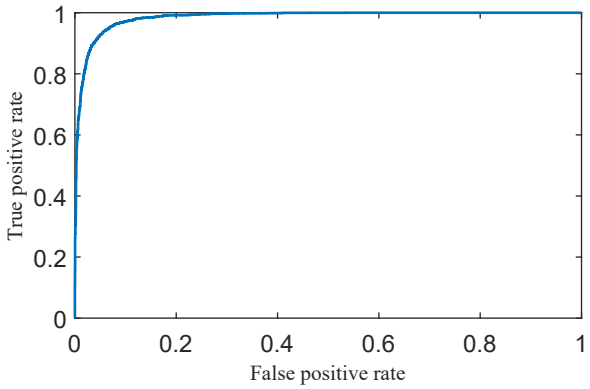


Figure 5.8: The ROC curve for the proposed CNN.

In Fig. 5.9, it shows the loss and performance curves during the training of the proposed CNN. Each iteration is corresponding to an update to the

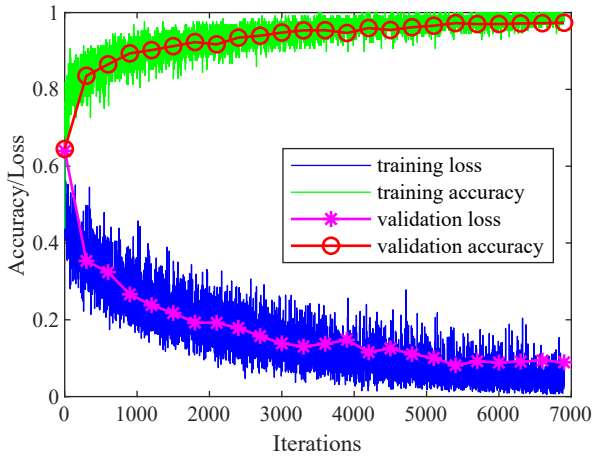


Figure 5.9: Loss and accuracy plots during the network training process.

network parameters based on the mini-batch gradients. The training accuracy and loss are calculated on each individual mini-batch. The validation accuracy and loss are calculated on the validation dataset at the end of every 300 iterations. The network training stops when the validation accuracy did not increase in the previous 5 validations, and it takes approximately 19 minutes. In Fig. 5.9, the training and validation accuracy curves are generally consistent during the network training process. The small divergence between the two at the end of the training, however, it may be related to the emerging of overfitting. Similar trend is also observed from the training and validation loss curves.

The 16 filters of the first convolutional layer are shown in Fig. 5.10. In Fig. 5.10, row 1-2 shows the two channels of the first 8 filters, and row 3-4 are the two channels of the other 8 filters. It is observed that both high and low frequency patterns are captured by using this group of filters. The similarity of the two filter channels may indicate that these filters could capture the symmetric properties of the input patch pair.

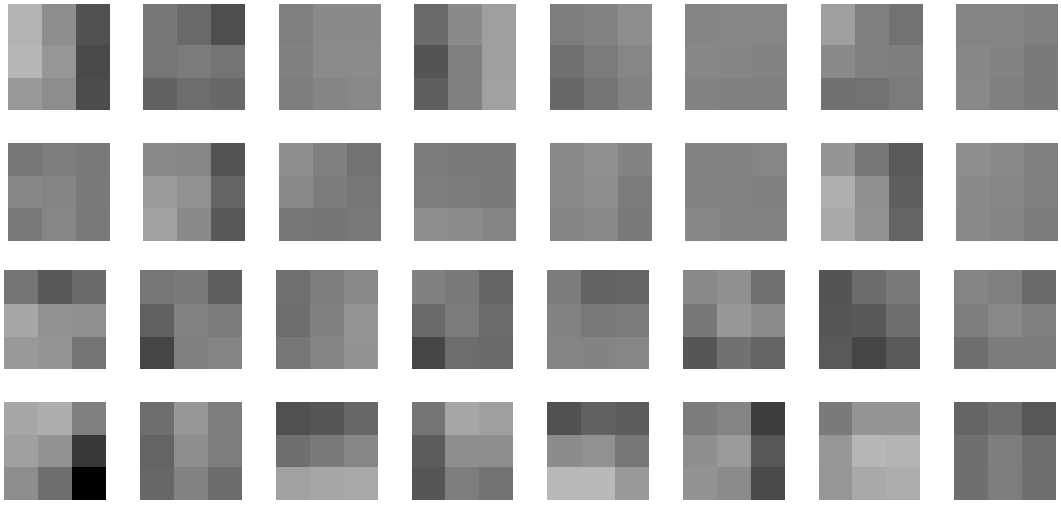


Figure 5.10: The first layer filters.

5.4.3 Evaluation of the System

To evaluate the FCD detection performance of the proposed system, we use leave-one-out cross validation strategy that all images but one are used for training the CNN model, and the left image is tested. This procedure is repeated until all images are evaluated. The FCD detection performance is measured using sensitivity, specificity, and accuracy.

To evaluate the FCD segmentation performance of the proposed system, the detected FCD patches are mapped to the inflated cortical surfaces (obtained using the FreeSurfer software [49]) and the dice coefficient is calculated using the following equation.

$$Dice = \frac{2(A \cap B)}{A \cup B} \tag{5.9}$$

where $A \cap B$ is the number of vertices common to the detected (A) and the ground-truth (B) FCD surface, and $A \cup B$ is the number of vertices in either A or B.

Table 5.3 shows the confusion matrix of the FCD detection. Nine of the ten FCD images are detected correctly using the proposed system. For non-

FCD images, 85% of them are correctly identified, and all of the three false detections are from the HFH dataset that has a larger slice thickness than the IXI dataset. It indicates the cortical structures may have some unnoticeable distortions due to either the effect of the disease (TLE) or the interpolation artifacts in image preprocessing. It is possible to increase the specificity by using higher resolution images.

Table 5.3: Confusion matrix for FCD detection with 30 MR images.

		Prediction		
		FCD	Non-FCD	Total
Actual	FCD	9	1	10
	Non-FCD	3	17	20
	Total	12	18	30

Fig. 5.11 shows an example of correct detections with a 3D view of the lesion on the inflated cortical surface. In Fig. 5.11a, it is an axial slice with marked FCD lesion along the red line. Patch extraction and classification results of this slice are shown in Fig. 5.11b and Fig. 5.11c, respectively. The magenta-color numbers in Fig. 5.11c are the predicted probabilities of these patches being FCD, and the high values at FCD regions illustrate the effectiveness of the proposed classification model. After post-processing, the FCD patches are mapped to the inflated cortical surface (obtained using FreeSurfer) so that the predicted FCD region could be viewed in 3D as the yellow region in Fig. 5.11d.

Fig. 5.12 shows the FCD image that is not successfully detected by the proposed technique. It is observed that although the lesion is subtle, the proposed CNN could classify the FCD patch with a high probability (0.99). But the post-processing stage mistakenly removed the FCD patch, resulting in the miss detection. To increase the sensitivity of the FCD detection, a smaller threshold (τ)) could be chosen, but it in turn will decrease the specificity.

Table 5.4 shows a performance comparison between the proposed technique

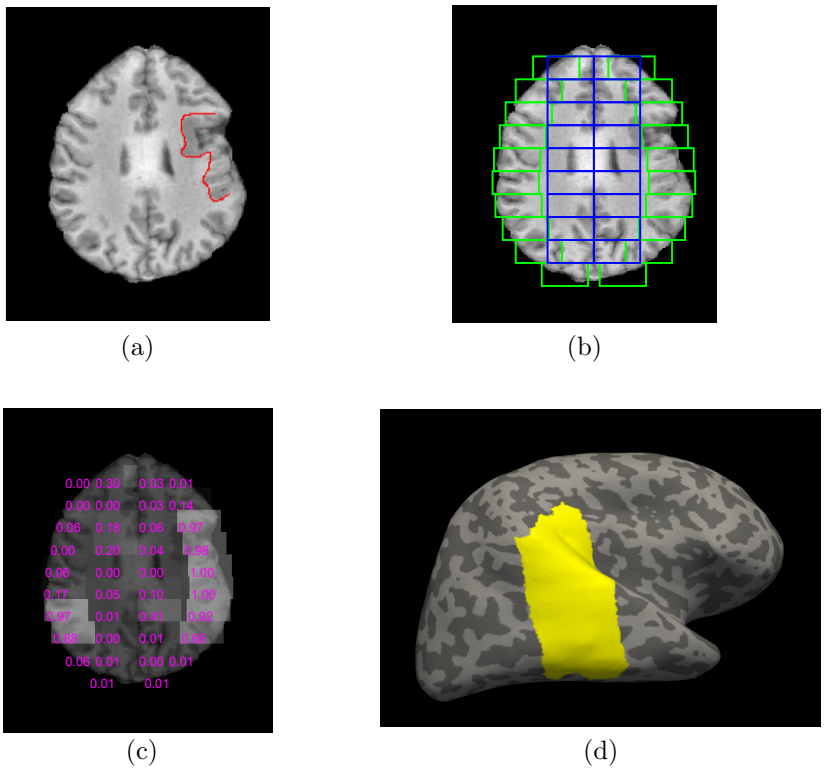
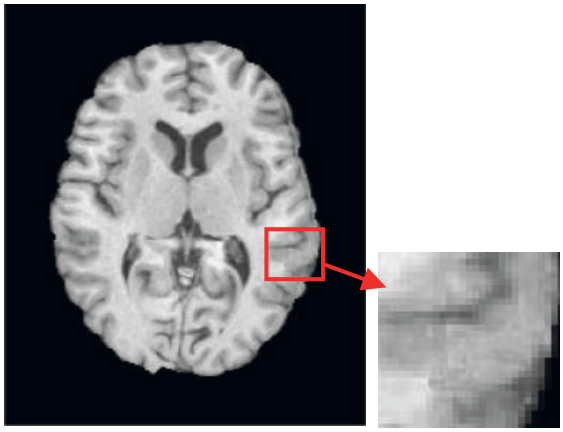


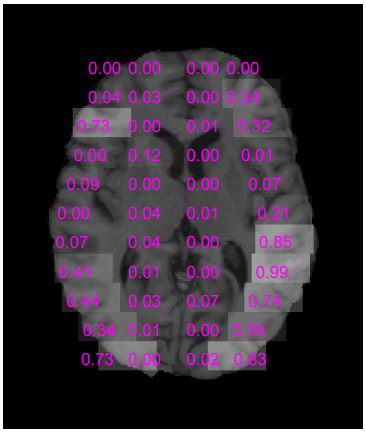
Figure 5.11: An example of FCD detection. (a) an axial slice with FCD lesion labeled in red, (b) patch extraction results, (c) classification results and numbers stand for probabilities of being FCD patches, (d) detection results mapped onto the inflated cortical surface (shown in yellow).

with three existing techniques using the images in this work. For the VBM technique [118], the optimal z-score threshold (of 3.5) is chosen to maximize the detection accuracy based on the GMC maps. The two SBC techniques [3, 82] are supervised classification techniques, and they are evaluated using the leave-one-out strategy. It is observed that the proposed technique provides a detection accuracy of 0.88, which is a 0.15 increase over the state of the art at 0.73. The specificity and sensitivity are also observed superior to the other three techniques. It indicates that the proposed deep learning technique may be a valuable tool for FCD detection and patient evaluation.

Table 5.5 compares the segmentation accuracy of the correctly detected FCD regions between the proposed technique and two SBC-based techniques



(a)



(b)

Figure 5.12: The missed FCD detection. (a) an axial slice with the enlarged FCD lesion, (b) patch probabilities (in magenta) of being FCD.

Table 5.4: Performance comparison with existing techniques for FCD detection. Spe.: specificity, sen.: sensitivity, acc: accuracy

Paper	Method	Spe.	Sen.	Acc.
Pail, 2012 [118]	VBM	0.60	0.70	0.66
Ahmed, 2015 [3]	SBC	0.75	0.60	0.70
Jin, 2018 [82]	SBC	0.75	0.70	0.73
Proposed	CNN	**0.85**	**0.90**	**0.88**

[3, 82] using surface mapping and dice coefficient. After FCD detection, we first reconstructed the cortical surface of each image using the FreeSurfer software, and then mapped the detection results to this surface. The dice coefficient is

calculated using Eq. 5.9 for each FCD lesion. The mean dice coefficients are reported in Table 5.5. It is observed that the proposed technique can provide a higher segmentation accuracy.

Table 5.5: Performance comparison with existing techniques for FCD segmentation.

Technique	Dice
Ahmed, 2015 [3]	0.43
Jin, 2018 [82]	0.55
Proposed	**0.78**

5.5 Conclusion

In this chapter, an automated technique is proposed for FCD detection in T1-weighted MR images using a deep convolutional neural network. The proposed technique first performs preprocessing to align the input image with a standard brain atlas. The cortical patches are then extracted on axial slices. Each patch is paired and fed to a CNN classifier with 5 convolutional layers, 1 max pooling layer, and 2 fully-connected layers. Finally, the post-processing stage removes noises and correct missed detections. Experimental results show a superior performance (90% sensitivity) of the proposed technique compared to the state of the art (70% sensitivity).

Chapter 6

Conclusions and Future Work

In recent years, computer-aided techniques of image analysis and lesion detection are attracting more and more attention from both academic researchers and medical professionals. These techniques are playing increasingly important roles in assisting neuroradiologists in their routine diagnosis and treatment planning. However, due to the variation and complexity of the human brain, much effort is still required to develop more robust computer-aided techniques for brain MRI image analysis and lesion detection. This thesis has developed several CAD techniques for the epileptogenic lesions, including cavernous malformation, mesial temporal sclerosis, and focal cortical dysplasia. This chapter concludes the thesis and discusses some future research works.

6.1 Conclusions

In Chapter 1, we presented a brief introduction to epilepsy and its diagnosis using MRI. The three most common epileptogenic lesions (*i.e.* CM, MTS, and FCD) were introduced, and the research topics of this thesis were presented. In Chapter 2, a brief review of related medical image analysis techniques was presented, followed by an overview of some existing CAD systems for different brain diseases. The main contributions of this thesis were presented in Chapters 3 to 5.

Chapter 3 presented the proposed technique for the automated detection of the cavernous malformation in T2-weighted axial plane MRI images. The brain area is first extracted using atlas registration and an active contour model. Template matching is then performed to obtain the candidate cavernoma regions. Texture, the histogram of oriented gradients and local binary pattern features of each candidate region are calculated, and principal component analysis is applied to reduce the feature dimensionality. Support vector machines are finally used to classify each region into cavernoma or non-cavernoma. The proposed technique was evaluated using 9 MR images of patients with a total of 23 cavernous malformations. The experimental results show a superior performance in cavernoma detection compared to existing techniques.

Chapter 4 presented the proposed technique for the automated detection of mesial temporal sclerosis based on the hippocampus and CSF features in MRI images. It includes brain segmentation and hippocampus extraction followed by calculating features of both hippocampus and its surrounding cerebrospinal fluid. Support vector machines are then applied to the generated features to identify patients with MTS from those without MTS. The proposed technique is developed and evaluated on a data set comprising 15 normal controls, 18 left and 18 right MTS patients. Experimental results show that subjects are correctly classified using the proposed classifiers with an accuracy of 0.94 for both left and right MTS detection.

Chapter 5 presented the proposed technique for the automated detection of focal cortical dysplasia using MRI images and deep learning. The input MRI image is first preprocessed to correct the bias field, normalize intensities, align with a standard atlas, and strip the non-brain tissues. All cortical patches are then extracted on each axial slice, and these patches are classified into FCD and non-FCD using a deep convolutional neural network (CNN) with five convolutional layers, a max pooling layer, and two fully-connected layers.

Finally, the false and missed classifications are corrected in the post-processing stage. The proposed technique is evaluated using images of 10 patients with FCD and 20 controls. The proposed CNN shows a superior performance in classifying cortical image patches compared with multiple CNN architectures. For the system-level evaluation, nine of the ten FCD images are successfully detected, and 85% of the non-FCD images are correctly identified. Overall, the proposed CNN-based technique improves the detection of FCD.

6.2 Resulting Publications

Based on the work of this thesis, the following papers have been published in refereed journals and conferences.

Journals

[1] H. Wang, SN. Ahmed and M. Mandal, Computer-aided diagnosis of cavernous malformations in brain MR images, *Computerized Medical Imaging and Graphics*, vol. 66, pp. 115-123, 2018.

[2] H. Wang, SN. Ahmed and M. Mandal, Computer-aided detection of mesial temporal sclerosis based on hippocampus and cerebrospinal fluid features in MR images, *Biocybernetics and Biomedical Engineering*, vol. 39, pp. 122-132, 2019.

[3] H. Wang, SN. Ahmed and M. Mandal, Automated detection of focal cortical dysplasia using a deep convolutional neural network, submitted, *Computerized Medical Imaging and Graphics*, 2018.

Conferences

[1] H. Wang, SN. Ahmed and M. Mandal, Efficient detection of mesial temporal sclerosis using hippocampus and CSF features in MRI images,

IEEE EMBS International Conference on Biomedical Health Informatics, Las Vegas, NV, 2018, pp. 178-181

[2] H. Wang, SN. Ahmed and M. Mandal, Automated detection of mesial temporal sclerosis based on multiple hippocampal features in T1-weighted MRI Images, *18th Annual Alberta Biomedical Engineering Conference*, Banff, AB, 2017, pp. 7

[3] H. Wang, SN. Ahmed, H. Xu and M. Mandal, Automated detection of cavernous malformations in brain MRI images, *8th IEEE EMBS Conference on Neural Engineering*, Shanghai, 2017, pp. 17-20.

[4] H. Wang, H. Xu, SN. Ahmed and M. Mandal, Computer aided detection of cavernous malformation in T2-weighted brain MR images, *IEEE Healthcare Innovation Point-Of-Care Technologies Conference*, Cancun, 2016, pp. 101-104.

6.3 Future Work

In this thesis, computer-aided detection techniques have been developed for the epileptogenic lesions. However, there are still many works that should be done to improve the completeness of a CAD system for epilepsy. A few possible future research works are introduced in this section.

6.3.1 More Epileptogenic Lesions

Other than cavernous malformation, mesial temporal sclerosis, and focal cortical dysplasia, there are still some other causes of intractable epilepsy, such as heterotopia, polymicrogyria, gliosis, and epileptogenic tumors. A complete CAD system for epilepsy should be able to screen as many epileptogenic lesions as possible. The techniques developed in this thesis could be extended

easily to detect these lesions. Future research work could be focused on developing detection techniques for more epileptogenic lesions, thus increasing the screening capability of the proposed CAD system for epilepsy.

6.3.2 Deep Learning in Epilepsy

In recent years, deep learning techniques, in particular, convolutional neural networks (CNN), have been used extensively in image classification and segmentation problems since they could learn optimal features automatically [99]. Some successful applications of deep learning in brain image analysis are brain extraction [90], brain segmentation [26], tumor segmentation [60], Alzheimer's disease prediction [70], tumor grading [119], microbleed detection [37], lacune detection [55], etc. In these applications, the deep learning techniques have already shown superior performance over the traditional techniques based on manually-crafted features.

The detection for the epileptogenic lesions in MRI is challenging due to the following reasons: a lack of labeled images, class imbalance, lesion complexity and variability in location, morphometry, and texture. In this thesis, we developed a CNN architecture for FCD that showed promising performance. Although several techniques have been proposed for the detection of epileptogenic lesions, all of these techniques are based on manually-crafted features (either voxel- or surface-based) [2, 71, 144, 163], and there is still room to improve the performance. Therefore, it is necessary to use deep learning algorithms in the computational analysis for epileptogenic lesions, and it is expected that the deep learning algorithms will increase the detection performance for epileptogenic lesions by automatically learning optimal lesion features in MRI.

6.3.3 Multiple Imaging Modalities

Currently, the computational analysis for epilepsy images is generally based on structural MRI. The 3D T1-weighted MR images are commonly used in the detection of mesial temporal sclerosis [93, 129], and focal cortical dysplasia [3, 163] for its high resolution and gray-white matter contrast. Few previous techniques have combined images of more than one imaging modalities in the epilepsy image analysis. Diffusion-weighted MRI generates images based on the diffusion property of water molecules [31], and it could increase the contrast of gray-white matter and the sensitivity to changes in gray matter [44]. Positron emission tomography (PET) uses radioactive tracers to observe the metabolic process and aid the diagnosis of lesions [12]. Some studies have shown that the co-registration of PET and MRI images improves the detection of cortical dysplasia [132]. Future research could be focused on combining features of multiple imaging modalities to increase the detection and segmentation performance for the epileptogenic lesion.

6.3.4 Accurate Segmentation of Epileptogenic Lesions

It is reported that approximately 80% of epilepsy patients are rendered seizure free after complete resection of epileptogenic lesions, while this number is only 20% - 50% when incomplete resection is performed due to lack of imaging findings [100]. It is important to segment epileptogenic lesions accurately and completely in MRI because of this correlation between complete resection and good surgical outcome. However, delineating epileptogenic lesions from healthy brain tissue is a difficult task [73], and very few studies have shown good segmentation results. Further research could be conducted on improving the accuracy of segmenting epileptogenic lesions in brain images.

References

[1] Understanding MRI gradients & slew rates. `https://www.amberusa.com/blog/understanding-mri-gradients-and-slew-rates`. Accessed: 2018-09-03.

[2] S. Adler, K. Wagstyl, R. Gunny, L. Ronan, D. Carmichael, J. H. Cross, et al. Novel surface features for automated detection of focal cortical dysplasias in paediatric epilepsy. *NeuroImage: Clinical*, 14:18–27, 2017.

[3] B. Ahmed, C. E. Brodley, K. E. Blackmon, R. Kuzniecky, G. Barash, C. Carlson, et al. Cortical feature analysis and machine learning improves detection of "MRI–negative" focal cortical dysplasia. *Epilepsy & Behavior*, 48:21–28, 2015.

[4] P. Aljabar, R. Heckemann, A. Hammers, J. Hajnal, and D. Rueckert. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *NeuroImage*, 46(3):726 – 738, 2009.

[5] R. D. Ambrosini, P. Wang, and W. G. O'Dell. Computer-aided detection of metastatic brain tumors using automated three-dimensional template matching. *Journal of Magnetic Resonance Imaging*, 31(1):85–93, 2010.

[6] V. Anitha and S. Murugavalli. Brain tumour classification using two-tier classifier with adaptive segmentation technique. *IET Computer Vision*, 10(1):9–17, 2016.

[7] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou. Lung pattern classification for interstitial lung dis-

eases using a deep convolutional neural network. *IEEE Transactions on Medical Imaging*, 35(5):1207–1216, May 2016.

[8] H. Arimura, T. Magome, Y. Yamashita, and D. Yamamoto. Computer-aided diagnosis systems for brain diseases in magnetic resonance images. *Algorithms*, 2(3):925–952, 2009.

[9] D. Armstrong. The neuropathology of temporal lobe epilepsy. *Journal of Neuropathology and Experimental Neurology*, 52(5):433–443, Sept. 1993.

[10] J. Ashburner. Computational anatomy with the SPM software. *Magnetic Resonance Imaging*, 27(8):1163–1174, 2009.

[11] J. Ashburner and K. J. Friston. Unified segmentation. *NeuroImage*, 26(3):839–851, 2005.

[12] D. L. Bailey, M. N. Maisey, D. W. Townsend, and P. E. Valk. *Positron emission tomography*. Springer, 2005.

[13] S. Batra, D. Lin, et al. Cavernous malformations: natural history, diagnosis and treatment. *Nature Reviews Neurology*, 5(12):659–670, Dec. 2009.

[14] S. Bauer, L. Nolte, and M. Reyes. Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 354–361, Toronto, Canada, Sept. 2011.

[15] S. Bauer, R. Wiest, L. Nolte, and M. Reyes. A survey of MRI-based medical image analysis for brain tumor studies. *Physics in medicine and biology*, 58(13):97–129, 2013.

[16] I. Beheshti, H. Demirel, et al. Structural MRI-based detection of Alzheimer's disease using feature ranking and classification error. *Computer Methods and Programs in Biomedicine*, 137:177–193, 2016.

[17] P. Besson, N. Bernasconi, O. Colliot, A. Evans, and A. Bernasconi. Surface-based texture and morphological analysis detects subtle cortical dysplasia. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 5241 of *Lecture Notes in Computer Science*, pages 645–652. Springer, Springer Berlin Heidelberg, 2008.

[18] M. Bland. *An introduction to medical statistics*. Oxford University Press (UK), 2015.

[19] L. Bonilha, J. J. Halford, C. Rorden, D. R. Roberts, Z. Rumboldt, and M. A. Eckert. Automated MRI analysis for identification of hippocampal atrophy in temporal lobe epilepsy. *Epilepsia*, 50(2):228–233, 2009.

[20] B. Braga, C. L. Yasuda, and F. Cendes. White matter atrophy in patients with mesial temporal lobe epilepsy: voxel-based morphometry analysis of T1-and T2-weighted MR images. *Radiology research and practice*, 2012:1–8, 2012.

[21] T. R. Browne and G. L. Holmes. *Handbook of epilepsy*. Jones & Bartlett Learning, 2008.

[22] D. L. Camacho and M. Castillo. MR imaging of temporal lobe epilepsy. *Seminars in Ultrasound, CT and MRI*, 28(6):424–436, 2007.

[23] D. Cantor-Rivera, A. R. Khan, M. Goubran, S. M. Mirsattari, and T. M. Peters. Detection of temporal lobe epilepsy using support vector machines in multi-parametric quantitative MR imaging. *Computerized Medical Imaging and Graphics*, 41:14–28, 2015.

[24] N. C. G. Centre. *The epilepsies: the diagnosis and management of the epilepsies in adults and children in primary and secondary care*. National Institute for Health and Clinical Excellence, 2012.

[25] Centre for the Developing Brain, Imperial College London. IXI dataset.

[26] H. Chen, Q. Dou, L. Yu, and P.-A. Heng. Voxresnet: Deep voxelwise residual networks for volumetric brain segmentation. *arXiv preprint arXiv:1608.05895*, 2016.

[27] D. Chicco. Ten quick tips for machine learning in computational biology. *BioData Mining*, 10(1):35, Dec. 2017.

[28] G. B. Coleman and H. C. Andrews. Image segmentation by clustering. *Proceedings of the IEEE*, 67(5):773–785, 1979.

[29] O. Colliot, N. Bernasconi, N. Khalili, S. B. Antel, V. Naessens, and A. Bernasconi. Individual voxel-based analysis of gray matter in focal cortical dysplasia. *Neuroimage*, 29(1):162–171, 2006.

[30] O. Colliot and et al. Discrimination between Alzheimer disease, mild cognitive impairment, and normal aging by using automated segmentation of the hippocampus. *Radiology*, 248(1):194–201, 2008.

[31] N. Colombo, N. Salamon, C. Raybaud, Ç. Özkara, and A. J. Barkovich. Imaging of malformations of cortical development. *Epileptic Disorders*, 11(3):194–205, 2009.

[32] C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[33] R. Cuingnet, C. Rosso, S. Lehéricy, D. Dormont, H. Benali, Y. Samson, and O. Colliot. Spatially regularized SVM for the detection of brain areas associated with stroke outcome. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 316–323. Springer, 2010.

[34] B. V. Dasarathy. *Nearest neighbor: pattern classification technique*. IEEE Computer Society, 1990.

[35] L. C. De Souza, M. Chupin, F. Lamari, C. Jardel, D. Leclercq, O. Colliot, S. Lehéricy, B. Dubois, and M. Sarazin. CSF tau markers are correlated

with hippocampal volume in Alzheimer's disease. *Neurobiology of Aging*, 33(7):1253–1257, 2012.

[36] S. R. Deans. *The Radon transform and some of its applications*. Courier Corporation, 2007.

[37] Q. Dou, H. Chen, L. Yu, L. Zhao, J. Qin, D. Wang, V. C. Mok, L. Shi, and P.-A. Heng. Automatic detection of cerebral microbleeds from mr images via 3d convolutional neural networks. *IEEE transactions on medical imaging*, 35(5):1182–1195, 2016.

[38] T. Dziedzic, P. Kunert, E. Matyja, K. Ziora-Jakutowicz, A. Sidoti, and A. Marchel. Familial cerebral cavernous malformation. *Folia Neuropathol*, 50(2):152–8, 2012.

[39] M. Eadie. Shortcomings in the current treatment of epilepsy. *Expert review of neurotherapeutics*, 12(12):1419–1427, 2012.

[40] E. A. El-Dahshan, H. Mohsen, et al. Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm. *Expert systems with Applications*, 41(11):5526–5545, 2014.

[41] M. J. England, C. T. Liverman, A. M. Schultz, and L. M. Strawbridge. Epilepsy across the spectrum: promoting health and understanding.: a summary of the institute of medicine report. *Epilepsy & Behavior*, 25(2):266–276, 2012.

[42] N. Farid, H. M. Girard, N. Kemmotsu, M. E. Smith, S. W. Magda, W. Y. Lim, R. R. Lee, and C. R. McDonald. Temporal lobe epilepsy: quantitative MR volumetry in detection of hippocampal atrophy. *Radiology*, 264(2):542–550, 2012.

[43] R. Farjam, H. A. Parmar, et al. An approach for computer-aided detection of brain metastases in post-Gd T1-W MRI. *Magnetic resonance imaging*, 30(6):824–836, 2012.

[44] K. W. Feindel. Can we develop pathology-specific mri contrast for "mr-negative" epilepsy? *Epilepsia*, 54:71–74, 2013.

[45] R. S. Fisher et al. ILAE official report: a practical clinical definition of epilepsy. *Epilepsia*, 55(4):475–482, 2014.

[46] N. K. Focke, P. J. Thompson, and J. S. Duncan. Correlation of cognitive functions with voxel-based morphometry in patients with hippocampal sclerosis. *Epilepsy & Behavior*, 12(3):472–476, 2008.

[47] N. K. Focke, M. Yogarajah, M. R. Symms, O. Gruber, W. Paulus, and J. S. Duncan. Automated MR image classification in temporal lobe epilepsy. *Neuroimage*, 59:356–362, 2012.

[48] V. S. Fonov, A. C. Evans, et al. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47, Supplement 1(47):102, 2009.

[49] A. A. M. C. for Biomedical Imaging". FreeSurfer software suite. Online: https://surfer.nmr.mgh.harvard.edu/.

[50] N. D. Forkert, T. Illies, E. Goebell, J. Fiehler, D. Säring, and H. Handels. Computer-aided nidus segmentation and angiographic characterization of arteriovenous malformations. *International Journal of Computer Assisted Radiology and Surgery*, 8(5):775–786, Sept. 2013.

[51] S. Free, P. Bergin, D. Fish, M. Cook, S. Shorvon, and J. Stevens. Methods for normalization of hippocampal volumes measured with MR. *American Journal of Neuroradiology*, 16(4):637–643, 1995.

[52] D. Fuerst, J. Shah, et al. Hippocampal sclerosis is a progressive disorder: a longitudinal volumetric MRI study. *Annals of neurology*, 53(3):413–416, 2003.

[53] A. Gajofatto, M. Calabrese, M. D. Benedetti, and S. Monaco. Clinical, MRI, and CSF markers of disability progression in multiple sclerosis. *Disease Markers*, 35(6):687–699, 2013.

[54] J. Gao and M. Xie. Skull-stripping MR brain images using anisotropic diffusion filtering and morphological processing. In *International Symposium on Computer Network and Multimedia Technology*, pages 1–4, Wuhan, China, Dec. 2009.

[55] M. Ghafoorian, N. Karssemeijer, T. Heskes, M. Bergkamp, J. Wissink, J. Obels, K. Keizer, F.-E. de Leeuw, B. van Ginneken, E. Marchiori, et al. Deep multi-scale location-aware 3d convolutional neural networks for automated detection of lacunes of presumed vascular origin. *NeuroImage: Clinical*, 14:391–399, 2017.

[56] X. Glorot, A. Bordes, and Y. Bengio. Deep sparse rectifier neural networks. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 315–323, 2011.

[57] J. A. González Torre, Á. J. Cruz-Gómez, A. Belenguer, C. Sanchis-Segura, C. Ávila, and C. Forn. Hippocampal dysfunction is associated with memory impairment in multiple sclerosis: a volumetric and functional connectivity study. *Multiple Sclerosis Journal*, 23(14):1854–1863, 2017.

[58] L. Guo, L. Zhao, Y. Wu, Y. Li, G. Xu, and Q. Yan. Tumor detection in MR images using one-class immune feature weighted SVMs. *IEEE Transactions on Magnetics*, 47(10):3849–3852, 2011.

[59] A. Hammers, R. Heckemann, M. J. Koepp, J. S. Duncan, J. V. Hajnal, D. Rueckert, and P. Aljabar. Automatic detection and quantification of hippocampal atrophy on MRI in temporal lobe epilepsy: a proof-of-principle study. *Neuroimage*, 36(1):38–47, 2007.

[60] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle. Brain tumor segmentation with deep neural networks. *Medical image analysis*, 35:18–31, 2017.

[61] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image

recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[62] R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage*, 33(1):115–126, 2006.

[63] P. Hellier. Consistent intensity correction of MR images. In *IEEE International Conference on Image Processing*, volume 1, page 1109, 2003.

[64] Henry Ford Hospital. Temporal lobe epilepsy multiple modality data, 2017.

[65] A. E. Hoerl and R. W. Kennard. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

[66] R. E. Hogan, R. D. Bucholz, and S. Joshi. Hippocampal deformation-based shape analysis in epilepsy and unilateral mesial temporal sclerosis. *Epilepsia*, 44(6):800–806, 2003.

[67] A. I. Holodny, A. E. George, J. Golomb, M. de Leon, and A. J. Kalnin. The perihippocampal fissures: normal anatomy and disease states. *Radiographics*, 18(3):653–665, 1998.

[68] S.-J. Hong, H. Kim, D. Schrader, N. Bernasconi, B. C. Bernhardt, and A. Bernasconi. Automated detection of cortical dysplasia type II in MRI-negative epilepsy. *Neurology*, 83(1):48–55, 2014.

[69] M.-P. Hosseini, M.-R. Nazem-Zadeh, D. Pompili, K. Jafari-Khouzani, K. Elisevich, and H. Soltanian-Zadeh. Comparative performance evaluation of automated segmentation methods of hippocampus from magnetic resonance images of temporal lobe epilepsy patients. *Medical Physics*, 43(1):538–553, 2016.

[70] E. Hosseini-Asl, G. Gimel'farb, and A. El-Baz. Alzheimer's disease diagnostics by a deeply supervised adaptable 3d convolutional network. *arXiv preprint arXiv:1607.00556*, 2016.

[71] P. M. House, M. Lanz, et al. Comparison of morphometric analysis based on T1- and T2-weighted {MRI} data for visualization of focal cortical dysplasia. *Epilepsy Research*, 106(3):403–409, 2013.

[72] J. Hua, Z. Xiong, J. Lowey, E. Suh, and E. R. Dougherty. Optimal number of features as a function of sample size for various classification rules. *Bioinformatics*, 21(8):1509–1515, 2004.

[73] H. Huppertz, C. Grimm, et al. Enhanced visualization of blurred gray–white matter junctions in focal cortical dysplasia by voxel–based 3D MRI analysis. *Epilepsy Research*, 67:35–50, 2005.

[74] S. Ioffe and C. Szegedy. Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv*, 2015.

[75] G. D. Jackson, S. F. Berkovic, J. S. Duncan, and A. Connelly. Optimizing the diagnosis of hippocampal sclerosis using MR imaging. *American journal of neuroradiology*, 14(3):753–762, 1993.

[76] K. Jafari-Khouzani, K. Elisevich, K. C. Karvelis, and H. Soltanian-Zadeh. Quantitative multi-compartmental SPECT image analysis for lateralization of temporal lobe epilepsy. *Epilepsy Research*, 95(1):35–50, 2011.

[77] K. Jafari-Khouzani, K. Elisevich, S. Patel, B. Smith, and H. Soltanian-Zadeh. FLAIR signal and texture analysis for lateralizing mesial temporal lobe epilepsy. *Neuroimage*, 49(2):1559–1571, 2010.

[78] F. Jäger, Y. Deuerling-Zheng, et al. A new method for MRI intensity standardization with application to lesion detection in the brain. *Procs*, 1010:269–276, 2006.

[79] R. Jenkins, N. C. Fox, A. M. Rossor, R. J. Harvey, and M. N. Rossor. Intracranial volume and alzheimer disease: evidence against the cerebral reserve hypothesis. *Archives of Neurology*, 57(2):220–224, 2000.

[80] M. Jenkinson, P. Bannister, M. Brady, and S. Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2):825–841, 2002.

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

[82] B. Jin, B. Krishnan, S. Adler, K. Wagstyl, W. Hu, S. Jones, I. Najm, A. Alexopoulos, K. Zhang, J. Zhang, et al. Automated detection of focal cortical dysplasia type II with surface-based magnetic resonance imaging postprocessing and machine learning. *Epilepsia*, 59(5):982–992, 2018.

[83] I. Jolliffe. *Principal component analysis*. Wiley Online Library, 2002.

[84] D. W. H. Jr. and S. Lemeshow. *Applied logistic regression*. Wiley, second edition, 2005.

[85] J. Kabat and P. Król. Focal cortical dysplasia–review. *Polish journal of radiology*, 77(2):35, 2012.

[86] P. Kalavathi and V. B. S. Prasath. Methods on skull stripping of MRI head scan images - a review. *Journal of Digital Imaging*, 29(3):365–379, 2016.

[87] W. T. Kerr, S. T. Nguyen, A. Y. Cho, E. P. Lau, D. H. Silverman, P. K. Douglas, N. M. Reddy, A. Anderson, J. Bramen, N. Salamon, et al. Computer-aided diagnosis and localization of lateralized temporal lobe epilepsy using interictal FDG-PET. *Frontiers in Neurology*, 4:31, 2013.

[88] N. E. A. Khalid, S. Ibrahim, and P. Haniff. MRI brain abnormalities segmentation using k-nearest neighbors(k-NN). *International Journal on Computer Science and Engineering*, 3(2):980–990, 2011.

[89] R. Khayati et al. A novel method for automatic determination of different stages of multiple sclerosis lesions in brain MR FLAIR images. *Computerized Medical Imaging and Graphics*, 32(2):124–133, 2008.

[90] J. Kleesiek, G. Urban, A. Hubert, D. Schwarz, K. Maier-Hein, M. Bendszus, and A. Biller. Deep mri brain extraction: a 3d convolutional neural network for skull stripping. *NeuroImage*, 129:460–469, 2016.

[91] P. Klenerman, J. W. Sander, and S. D. Shorvon. Mortality in patients with epilepsy: a study of patients in long term residential care. *Journal of Neurology, Neurosurgery & Psychiatry*, 56(2):149–152, 1993.

[92] S. Klöppel et al. Automatic classification of MR scans in Alzheimer's disease. *Brain*, 131(3):681–689, 2008.

[93] Z. Kohan and R. Azmi. Hippocampus shape analysis for temporal lobe epilepsy detection in magnetic resonance imaging. In *Medical Imaging: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 9788, page 97882, San Diego, California, United States, Feb. 2016.

[94] H. Kong, H. C. Akakin, and S. E. Sarma. A generalized Laplacian of Gaussian filter for blob detection and its applications. *IEEE transactions on cybernetics*, 43(6):1719–1733, 2013.

[95] S. Kotsiantis, D. Kanellopoulos, et al. Handling imbalanced datasets: a review. *GESTS International Transactions on Computer Science and Engineering*, 30(1):25–36, 2006.

[96] E. O. Kreyszig. *Advanced engineering mathematics*. Wiley, 1979.

[97] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[98] E. Kumlien, R. C. Doss, and J. R. Gates. Treatment outcome in patients with mesial temporal sclerosis. *Seizure*, 11(7):413–417, 2002.

[99] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *nature*, 521(7553):436, 2015.

[100] J. T. Lerner, N. Salamon, J. S. Hauptman, T. R. Velasco, M. Hemb, J. Y. Wu, R. Sankar, W. Donald Shields, J. Engel Jr, I. Fried, et al. Assessment and surgical outcomes for mild type I and severe type II cortical dysplasia: a critical review and the UCLA experience. *Epilepsia*, 50(6):1310–1335, 2009.

[101] M. Liu, L. Concha, C. Lebel, C. Beaulieu, and D. W. Gross. Mesial temporal sclerosis is linked with more widespread white matter changes in temporal lobe epilepsy. *NeuroImage: clinical*, 1(1):99–105, 2012.

[102] C. P. Loizou, V. Murray, M. S. Pattichis, I. Seimenis, M. Pantziaris, and C. S. Pattichis. Multiscale amplitude-modulation frequency-modulation texture analysis of multiple sclerosis in brain MRI images. *IEEE Transactions on Information Technology in Biomedicine*, 15(1):119–129, Jan. 2011.

[103] K. Machhale, H. B. Nandpuru, et al. MRI brain cancer classification using hybrid classifier (SVM-KNN). In *IEEE International Conference on Industrial Instrumentation and Control (ICIC)*, pages 60–65, 2015.

[104] A. Madabhushi and J. K. Udupa. Interplay between intensity standardization and inhomogeneity correction in MR image processing. *IEEE Transactions on Medical Imaging*, 24(5):561–576, 2005.

[105] K. Malmgren and M. Thom. Hippocampal sclerosis - origins and imaging. *Epilepsia*, 53(Suppl. 4):19–33, 2012.

[106] J. N. Mandrekar. Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology*, 5(9):1315–1316, 2010.

[107] J. N. Maraire et al. Intracranial cavernous malformations lesion behavior and management strategies. *Neurosurgery*, 37(4):591–605, Oct. 1995.

[108] A. Mayer and H. Greenspan. An adaptive mean-shift framework for MRI brain segmentation. *IEEE Transactions on Medical Imaging*, 28(8):1238–1250, 2009.

[109] A. Mechelli, C. J. Price, K. J. Friston, and J. Ashburner. Voxel-based morphometry of the human brain: methods and applications. *Current medical imaging reviews*, 1(2):105–113, 2005.

[110] R. Meier, S. Bauer, J. Slotboom, R. Wiest, and M. Reyes. Appearance- and context-sensitive features for brain tumor segmentation. *Proceedings of MICCAI BRATS Challenge*, pages 20–26, 2014.

[111] L. Mumoli, A. Labate, R. Vasta, A. Cherubini, E. Ferlazzo, U. Aguglia, A. Quattrone, and A. Gambardella. Detection of hippocampal atrophy in patients with temporal lobe epilepsy: a 3-Tesla MRI shape. *Epilepsy & Behavior*, 28(3):489–493, 2013.

[112] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning*, pages 807–814, 2010.

[113] A. P. Nanthagopal and R. Sukanesh. Wavelet statistical texture features-based segmentation and classification of brain computed tomography images. *IET Image Processing*, 7(1):25–32, 2013.

[114] B. H. Nayef, S. Abdullah, et al. Brain images application and supervised learning algorithms: a review. *Journal of Medical Sciences*, 14(3):108, 2014.

[115] L. G. Nyúl, J. K. Udupa, et al. On standardizing the MR image intensity scale. *image*, 1081, 1999.

[116] E. K. Oermann, A. Rubinsteyn, D. Ding, J. Mascitelli, R. M. Starke, J. B. Bederson, H. Kano, L. D. Lunsford, J. P. Sheehan, J. Hammerbacher, et al. Using a machine learning approach to predict outcomes after radiosurgery for cerebral arteriovenous malformations. *Scientific Reports*, 6:21161, 2016.

[117] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285–296):23–27, 1975.

[118] M. Pail, R. Mareček, M. Hermanová, B. Slaná, I. Tyrlíková, R. Kuba, and M. Brázdil. The role of voxel-based morphometry in the detection of cortical dysplasia within the temporal pole in patients with intractable mesial temporal lobe epilepsy. *Epilepsia*, 53(6):1004–1012, 2012.

[119] Y. Pan, W. Huang, Z. Lin, W. Zhu, J. Zhou, J. Wong, and Z. Ding. Brain tumor grading based on neural networks and convolutional neural networks. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 699–702. IEEE, 2015.

[120] B. Patenaude, S. M. Smith, D. N. Kennedy, and M. Jenkinson. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage*, 56(3):907–922, 2011.

[121] S. Pereira, A. Pinto, V. Alves, and C. A. Silva. Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging*, 35(5):1240–1251, May 2016.

[122] J. C. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in Large Margin Classifiers*, pages 61–74. MIT Press, 1999.

[123] J. C. Platt, N. Cristianini, and J. Shawe-Taylor. Large margin DAGs for multiclass classification. In *Advances in neural information processing systems*, pages 547–553, 2000.

[124] P. J. Porter, R. A. Willinsky, et al. Cerebral cavernous malformations: natural history and prognosis after clinical deterioration with or without hemorrhage. *Journal of neurosurgery*, 87(2):190–197, 1997.

[125] X. Qu, L. Platisa, et al. Estimating blur at the brain gray-white matter boundary for FCD detection in MRI. In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3321–3324, Aug. 2014.

[126] J. Rajan, K. Kannan, C. Kesavadas, and B. Thomas. Focal cortical dysplasia (FCD) lesion analysis with complex diffusion approach. *Computerized Medical Imaging and Graphics*, 33(7):553–558, 2009.

[127] R. Riccelli, N. Toschi, S. Nigro, A. Terracciano, and L. Passamonti. Surface-based morphometry reveals the neuroanatomical basis of the five-factor model of personality. *Social cognitive and affective neuroscience*, 12(4):671–684, 2017.

[128] C. Robert. Machine learning, a probabilistic perspective. *CHANCE*, 27(2):62–63, 2014.

[129] J. D. Rudie, J. B. Colby, and N. Salamon. Machine learning classification of mesial temporal sclerosis in epilepsy patients. *Epilepsy Research*, 117:63–69, 2015.

[130] C. RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29(3):162–173, June 1996.

[131] B. N. Saha, N. Ray, et al. Quick detection of brain tumors and edemas: a bounding box method using symmetry. *Computerized medical imaging and graphics*, 36(2):95–107, 2012.

[132] N. Salamon, J. Kung, S. Shaw, J. Koo, S. Koh, J. Wu, J. Lerner, R. Sankar, W. Shields, J. Engel, et al. FDG-PET/MRI coregistration improves detection of cortical dysplasia in patients with epilepsy. *Neurology*, 71(20):1594–1601, 2008.

[133] D. H. Salat, R. L. Buckner, A. Z. Snyder, D. N. Greve, R. S. Desikan, E. Busa, J. C. Morris, A. M. Dale, and B. Fischl. Thinning of the cerebral cortex in aging. *Cerebral cortex*, 14(7):721–730, 2004.

[134] D. Shattuck, S. Sandor-Leahy, K. Schaper, D. Rottenberg, and R. Leahy. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13(5):856–876, 2001.

[135] S. Shinnar. Febrile seizures and mesial temporal sclerosis. *Epilepsy Currents*, 3:115–118, July 2003.

[136] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv*, 2014.

[137] S. Smith. Fast robust automated brain extraction. *Human brain mapping*, 17(3):143–155, 2002.

[138] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23:208–219, 2004.

[139] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958, 2014.

[140] N. H. Stricker, H. H. Dodge, N. M. Dowling, S. D. Han, E. A. Erosheva, W. J. Jagust, et al. CSF biomarker associations with change in hippocampal volume and precuneus thickness: implications for the Alzheimer's pathological cascade. *Brain imaging and behavior*, 6(4):599–609, 2012.

[141] M. Suzuki, H. Hagino, S. Nohara, S.-Y. Zhou, Y. Kawasaki, T. Takahashi, M. Matsui, H. Seto, T. Ono, and M. Kurachi. Male-specific volume expansion of the human hippocampus during adolescence. *Cerebral Cortex*, 15(2):187–193, 2004.

[142] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, et al. Going deeper with convolutions. Cvpr, 2015.

[143] X. Tai, B. Bernhardt, M. Thom, P. Thompson, S. Baxendale, M. Koepp, and N. Bernasconi. Neurodegenerative processes in temporal lobe

epilepsy with hippocampal sclerosis: clinical, pathological and neuroimaging evidence. *Neuropathology and Applied Neurobiology*, 44(1):70–90, 2018.

[144] Y.-L. Tan, H. Kim, et al. Quantitative surface analysis of combined MRI and PET enhances detection of focal cortical dysplasias. *Neuroimage*, 166:10–18, 2018.

[145] F. H. Tomlinson et al. Angiographically occult vascular malformations a correlative study of features on magnetic resonance imaging and histological examination. *Neurosurgery*, 34(5):792–800, 1994.

[146] J. Trefnỳ and J. Matas. Extended set of local binary patterns for rapid object detection. In *Proceedings of the computer vision winter workshop*, pages 1–7, Nové Hrady, Czech Republic, Feb. 2010.

[147] U. Vovk, F. Pernus, and B. Likar. A review of methods for correction of intensity inhomogeneity in MRI. *IEEE transactions on medical imaging*, 26(3):405–421, 2007.

[148] J. Wagner, B. Weber, H. Urbach, C. E. Elger, and H.-J. Huppertz. Morphometric MRI analysis improves detection of focal cortical dysplasia type II. *Brain*, 134(10):2844–2854, 2011.

[149] H. Wang, S. N. Ahmed, and M. Mandal. Automated detection of mesial temporal sclerosis based on multiple hippocampal features in T1-weighted MRI images. In *8th Annual Alberta Biomedical Engineering Conference*, page 7, Nov. 2017.

[150] H. Wang, S. N. Ahmed, and M. Mandal. Automated detection of focal cortical dysplasia using a deep convolutional neural network. *Computerized Medical Imaging and Graphics (submitted)*, 2018.

[151] H. Wang, S. N. Ahmed, and M. Mandal. Computer-aided detection of mesial temporal sclerosis based on hippocampus and cerebrospinal

fuid features in MR images. *Biocybernetics and Biomedical Engineering (accepted)*, 2018.

[152] H. Wang, S. N. Ahmed, and M. Mandal. Computer-aided diagnosis of cavernous malformations in brain MR images. *Computerized Medical Imaging and Graphics*, 66:115–123, 2018.

[153] H. Wang, S. N. Ahmed, and M. Mandal. Efficient detection of mesial temporal sclerosis using hippocampus and CSF features in MRI images. In *IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, pages 178–181, Mar. 2018.

[154] H. Wang, S. N. Ahmed, H. Xu, and M. Mandal. Automated detection of cavernous malformations in brain MRI images. In *8th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 17–20. IEEE, 2017.

[155] H. Wang, H. Xu, S. N. Ahmed, and M. Mandal. Computer aided detection of cavernous malformation in T2-weighted brain MR images. In *IEEE Healthcare Innovation Point-Of-Care Technologies Conference (HI-POCT)*, pages 101–104. IEEE, 2016.

[156] H. Wang, H. Xu, et al. Computer aided detection of cavernous malformation in T2-weighted brain MR images. In *IEEE Healthcare Innovation Point-Of-Care Technologies Conference (HI-POCT)*, pages 101–104, Cancun, Mexico, Nov. 2016.

[157] S. H. Wang et al. Multiple sclerosis detection based on biorthogonal wavelet transform, RBF kernel principal component analysis, and logistic regression. *IEEE Access*, 4:7567–7576, 2016.

[158] X. Wang, T. X. Han, and S. Yan. An HOG-LBP human detector with partial occlusion handling. In *IEEE International Conference on Computer Vision*, pages 32–39, Kyoto, Japan, Sept. 2009.

[159] Y. Wang, Y. Fan, P. Bhatt, and C. Davatzikos. High-dimensional pattern regression using machine learning: from medical images to continuous clinical variables. *Neuroimage*, 50(4):1519–1535, 2010.

[160] W. Wein, S. Brunke, A. Khamene, et al. Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. *Medical image analysis*, 12(5):577–585, 2008.

[161] Wellcome Trust Centre for Neuroimaging at UCL. SPM - statistical parametric mapping. `http://www.fil.ion.ucl.ac.uk/spm/`, 2014.

[162] S. Wiebe, W. T. Blume, J. P. Girvin, and M. Eliasziw. A randomized, controlled rrial of surgery for temporal-lobe epilepsy. *New England Journal of Medicine*, 345(5):311–318, 2001.

[163] L. C. Wong-Kisiel, D. F. T. Quiroga, et al. Morphometric analysis on T1-weighted MRI complements visual MRI review in focal cortical dysplasia. *Epilepsy Research*, 140:184–191, 2018.

[164] C. Xu and J. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on image processing*, 7(3):359–369, 1998.

[165] Y. Yamashita, H. Arimura, and K. Tsuchiya. Computer-aided detection of ischemic lesions related to subcortical vascular dementia on magnetic resonance images. *Academic radiology*, 15(8):978–985, 2008.

[166] S. W. Yoon, H. K. Lee, J. H. Kim, and M. H. Lee. Medical endoscopic image segmentation using snakes. *IEICE Transactions on Information and Systems*, 87(3):785–789, 2004.

[167] Y. Yu, D. Guo, M. Lou, D. S. Liebeskind, and F. Scalzo. Prediction of hemorrhagic transformation severity in acute stroke from source perfusion MRI. *IEEE Transactions on Biomedical Engineering*, PP(99):1, 2017.

[168] J. Zhang, Y. Gao, Y. Gao, B. C. Munsell, and D. Shen. Detecting anatomical landmarks for fast Alzheimer's disease diagnosis. *IEEE Transactions on Medical Imaging*, 35(12):2524–2533, Dec. 2016.

[169] Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1):45–57, Jan. 2001.

[170] Y. Zhuge and J. K. Udupa. Intensity standardization simplifies brain MR image segmentation. *Computer vision and image understanding*, 113(10):1095–1103, 2009.

[171] A. P. Zijdenbos, R. Forghani, et al. Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. *IEEE transactions on medical imaging*, 21(10):1280–1291, 2002.