

Volumetric Data Classification: A Study Directed at 3-D  
Imagery

Thesis submitted in accordance with the requirements of  
the University of Liverpool for the degree of Doctor in Philosophy  
by  
Akadej Udomchaiporn

July 2016

# Declaration

A number of publications have arisen out of the work presented in this thesis; some of the later chapters are closely correlated to a number of these publications as follows:

1. *A. Udomchaiporn, F. Coenen, V. Sluming, and M. García-Fiñana. 3-D MRI Brain Scan Feature Classification Using an Oct-tree Representation. The 9<sup>th</sup> International Conference on Advanced Data Mining and Applications (ADMA 2013). pp. 229-240. Springer Berlin Heidelberg, 2013.* [122] This paper presented a procedure for the classification of specific VOI (the ventricles) found in MRI brain scan volumes using an oct-tree representation technique similar to that presented in Chapter 7. As suggested later in this thesis, Frequent Sub-graph Mining (FSM) was used to define a feature space. The paper also introduced preliminary versions of the Volume Growing and Bounding Box segmentation algorithms for isolating the lateral ventricles within the MRI data, updated versions of these algorithms are presented in Chapter 4. The dataset used for the evaluation was the “Epilepsy” dataset, which was also used with respect to the evaluations presented later in this thesis.
2. *A. Udomchaiporn, F. Coenen, V. Sluming, and M. García-Fiñana. 3-D MRI Brain Scan Classification of Epilepsy Versus Non-epilepsy. The 18<sup>th</sup> Annual Conference in Medical Image Understanding and Analysis (MIUA 2014). pp. 253-258. London, UK. 9-11 July 2014.* [123] This paper presented a 3-D classification method, for application to MRI brain scan volumes, founded on two Point Series representation techniques, techniques that are also presented later in this theses in Chapter 6. The two point series representations considered were: (i) Disc based and (ii) Spoke based. The evaluation presented was directed at identifying volumes that feature indicators of epilepsy against volumes that do not. The epilepsy data set was thus again used in this paper for evaluation purposes.
3. *A. Udomchaiporn, F. Coenen, V. Sluming, and M. García-Fiñana. 3-D MRI Brain Scan Classification Using A Point Series Based Representation. The 16<sup>th</sup> International Conference of Data Warehousing and Knowledge Discovery (DaWaK 2014). pp. 300-307. Munich, Germany. September 2-4, 2014.* [124] This paper

reconsidered the point series based MRI VOI representation but instead of using a Hough signature feature space based classification approach, as advocated in the previous publication, proposed the use of a KNN algorithm coupled with Warping Distance as the similarity measure. The approach also features later in this thesis in Chapter 6.

4. *A. Udomchaiporn, F. Coenen, V. Sluming, and M. García-Fiñana. 3-D Volume of Interest Based Image Classification. The 14<sup>th</sup> Pacific Rim International Conference on Artificial Intelligence (PRICAI 2016). Phuket, Thailand. August 22-26, 2016. [125]* This paper proposed all the three classification techniques: (i) Statistical Metrics, (ii) Point Series and (iii) Oct-tree and also considered augmenting the VOI data with meta data. The paper also proposed statistical comparison among the proposed techniques. The approaches feature later in this thesis in Chapter 5, 6, 7, and 8.

# Abstract

This thesis describes research work undertaken in the field of image mining (particularly medical image mining). More specifically, the research work is directed at 3-D image classification according to the nature of a particular Volume Of Interest (VOI) that appears across a given image set. In this thesis the term VOI Based Image Classification (VOIBIC) has been coined to describe this process. VOIBIC entails a number of challenges. The first is the identification and isolation of the VOIs. Two segmentation algorithms are thus proposed to extract a given VOI from an image set: (i) Volume Growing and (ii) Bounding Box. The second challenge that VOIBIC poses is, once the VOI have been identified, how best to represent the VOI so that classification can be effectively and efficiently conducted. Three approaches are considered. The first is founded on the idea of using statistical metrics, the Statistical Metrics based representation. This representation offers the advantage in that it is straightforward and, although not especially novel, provides a benchmark. The second proposed representation is founded on the concept of point series (curves) describing the perimeter of a VOI, the Point Series representation. Two variations of this representation are considered: (i) Spoke based and (ii) Disc based. The third proposed representation is founded on a Frequent Subgraph Mining (FSM) technique whereby the VOI is represented using an Oct-tree structure to which FSM can be applied. The identified frequent subtrees can then be used to define a feature vector representation compatible with many classifier model generation methods. The thesis also considers augmenting the VOI data with meta data, namely age and gender, and determining the effect this has on performance. The presented evaluation used two 3-D MRI brain scan data sets: (i) Epilepsy and (ii) Musicians. The VOI in this case were the lateral ventricles, a distinctive VOI in such MRI brain scan data. For evaluation purposes two scenarios are considered, distinguishing between: (i) epilepsy patients and healthy people and (ii) musicians and non-musicians. The results indicates that the Spoke based point series representation technique produced the best results with a recorded classification accuracy of up to 78.52% for the Epilepsy dataset and 84.91% for the Musician dataset.

# Acknowledgement

I would like to express my deepest gratitude to my first supervisor Prof.Frans Coenen who has been abundantly helpful and has offered invaluable assistance, support and guidance throughout my Ph.D. study which has made it one of the most unforgettable experiences ever. It is my great honour to have worked with him. I would also like to offer special thanks to my second and third supervisors Dr. Vanessa Sluming and Dr. Marta Van der Hoek. Without whose knowledge and assistance this research would not have been successful. I would like to say thank you to all the staff and friends in the Department of Computer Science at The University of Liverpool, and my friends from Thailand, that have always been there for me whenever needed. I would also like to convey thanks to the Royal Thai Government for providing me the opportunity and financial means to conduct my Ph.D. study. Finally, and most importantly, I would like to offer infinite thanks and love to my family, for their unconditional support for everything I wish to achieve; and without whom it would be impossible for me to have a chance at undertaking my Ph.D.

# Contents

<b>Declaration</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Acknowledgement</b>	<b>iv</b>
<b>Contents</b>	<b>viii</b>
<b>List of Figures</b>	<b>xii</b>
<b>List of Tables</b>	<b>xiv</b>
<b>Nomenclature</b>	<b>xiv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	2
1.2 Thesis Objectives . . . . .	3
1.3 Research Methodology . . . . .	5
1.4 Thesis Contributions . . . . .	6
1.5 Thesis Organisation . . . . .	7
1.6 Conclusion . . . . .	8
<b>2 Background and Related Work</b>	<b>9</b>
2.1 Application Domain . . . . .	10
2.1.1 Magnetic Resonance Imaging (MRI) . . . . .	10
2.1.2 Medical Conditions Relating to MRI Brain Scan . . . . .	12
2.2 Review of Image Segmentation . . . . .	13
2.2.1 Image Segmentation Using Thresholding . . . . .	13
Global Thresholding . . . . .	13
Local Thresholding . . . . .	14
Hysteresis Thresholding . . . . .	15
2.2.2 Image Segmentation Using Region Growing . . . . .	16
Growing by Value . . . . .	16

Adaptive Region Growing . . . . .	16
Adams Seeded Region Growing . . . . .	16
Non-Connected Region Growing . . . . .	17
Parameter-Free Region Growing . . . . .	17
2.2.3 Deformable Methods and Level Set Methods . . . . .	17
Deformable Models . . . . .	17
Level Sets . . . . .	17
2.2.4 Other Segmentation Methods . . . . .	18
Image Segmentation Using Fuzzy Correctedness . . . . .	18
Image Segmentation Using Watershed Algorithm . . . . .	18
Image Segmentation Using Bayesian Methods . . . . .	18
2.2.5 Image Segmentation in MRI Brain Scan Data . . . . .	19
2.3 Review of Image Representation . . . . .	20
2.3.1 Statistical Based Image Representation . . . . .	20
2.3.2 Discrete Wavelet Transform (DWT) Based Image Representation	22
2.3.3 Histogram Based Image Representation . . . . .	22
Simple Histograms . . . . .	22
Histograms of Oriented Gradients (HOGs) . . . . .	22
Histograms of Local Binary Patterns (LBPs) . . . . .	23
Histograms of Local Phase Quantisation (LPQ) . . . . .	23
2.3.4 Hough Transform . . . . .	23
2.3.5 Tree Based Image Representation . . . . .	24
2.4 Review of Selected Classification Mechanisms . . . . .	25
2.4.1 Decision Tree . . . . .	25
2.4.2 Support Vector Machine . . . . .	26
2.4.3 Bayesian Network . . . . .	28
2.4.4 Artificial Neural Network . . . . .	29
2.4.5 K-Nearest Neighbours . . . . .	31
2.5 Image Classification in MRI Brain Scan . . . . .	32
2.6 Evaluation Criteria . . . . .	32
2.7 Conclusion . . . . .	34
<b>3 Evaluation Datasets</b>	<b>35</b>
3.1 Epilepsy Dataset . . . . .	36
3.2 Musician Dataset . . . . .	38
3.3 Conclusion . . . . .	40
<b>4 Volume of Interest Identification</b>	<b>42</b>
4.1 Image Preprocessing . . . . .	44
4.1.1 Slice Capture and Registration . . . . .	44

4.1.2	Contrast Enhancement . . . . .	45
4.2	Image Segmentation . . . . .	47
4.2.1	3-D Segmentation Using The Volume Growing Algorithm . . . . .	47
4.2.2	3-D Segmentation Using The Bounding Box Algorithm . . . . .	50
4.3	Evaluation . . . . .	51
4.4	Conclusion . . . . .	61
<b>5</b>	<b>Volume of Interest Image Classification Using Statistical Metrics Based Representation</b>	<b>63</b>
5.1	Statistical Metrics Based Image Classification . . . . .	63
5.2	Evaluation . . . . .	64
5.2.1	Usage of Different Classification Models . . . . .	65
5.2.2	Usage of Meta Attributes . . . . .	66
5.2.3	Efficiency . . . . .	68
5.2.4	Comparison of Attributes . . . . .	71
5.3	Conclusion . . . . .	72
<b>6</b>	<b>Volume of Interest Image Classification Using Point Series Representation</b>	<b>73</b>
6.1	Point Series Image Representation . . . . .	74
6.1.1	Disc-based Representation Technique . . . . .	74
6.1.2	Spoke-based Representation Technique . . . . .	75
6.2	Framework for Image Classification Based on Point Series Representation	77
6.2.1	Feature Space Classification . . . . .	77
6.2.2	KNN Classification . . . . .	80
6.3	Evaluation . . . . .	83
6.3.1	Disc-Based versus Spoke-Based . . . . .	84
6.3.2	Best Value for $\theta$ . . . . .	87
6.3.3	Feature Space Classification versus KNN Classification . . . . .	88
6.3.4	Use of Meta Attributes . . . . .	89
6.4	Conclusion . . . . .	91
<b>7</b>	<b>Volume of Interest Image Classification Using Oct Tree Based Representation</b>	<b>92</b>
7.1	Oct-tree Image Representation . . . . .	93
7.2	Feature Extraction . . . . .	94
7.2.1	Graph Mining . . . . .	95
7.2.2	Frequent Subgraph Mining . . . . .	96
7.2.3	gSpan . . . . .	97
	The Average Total Weighting (ATW) Scheme . . . . .	99

The gSpan-ATW Algorithm . . . . .	100
7.2.4 Feature Vector Generation . . . . .	101
7.3 Evaluation . . . . .	101
7.3.1 Support Vector Machine classification versus Decision Tree Clas- sification . . . . .	102
7.3.2 Best Value of Minimum Support Threshold . . . . .	103
7.3.3 Use of Meta Attributes . . . . .	103
7.3.4 Run Time Efficiency . . . . .	105
7.4 Conclusion . . . . .	107
<b>8 Discussion</b>	<b>109</b>
8.1 Comparison in Terms of Earlier Evaluation Results . . . . .	109
8.2 Statistical Significance Comparison . . . . .	112
8.3 Comparison with Previous Work . . . . .	114
8.4 Conclusion . . . . .	115
<b>9 Conclusion</b>	<b>116</b>
9.1 Summary of the Proposed Approaches . . . . .	116
9.2 Main Findings and Contributions . . . . .	119
9.3 Potential Future work . . . . .	121
<b>Bibliography</b>	<b>133</b>
<b>Index</b>	<b>133</b>

# List of Figures

2.1	The main research areas of study related to the work present in this thesis: Magnetic Resonance Imaging (MRI), image representation and data mining . . . . .	9
2.2	Example of a MRI scanner . . . . .	11
2.3	Example of a MRI brain scanner with a “receiver frame” . . . . .	11
2.4	Example of a typical perceptron in ANN . . . . .	30
2.5	Example of confusion matrix . . . . .	33
3.1	Epilepsy MRI brain scan data grouped by gender . . . . .	36
3.2	Epilepsy MRI brain scan data grouped by age range . . . . .	37
3.3	The percentages of male and female subjects grouped by age, and the percentages of the Epilepsy and non-Epilepsy subjects grouped by age, for the Epilepsy MRI brain scan data . . . . .	37
3.4	Example of MRI brain scan from an Epilepsy patient . . . . .	38
3.5	Musician MRI brain scan data grouped by gender . . . . .	39
3.6	Musician MRI brain scan data grouped by age range . . . . .	39
3.7	The percentages of male and female subjects grouped by age, and the percentages of musicians and non-musicians grouped by age, for the Musician MRI brain scan data . . . . .	40
3.8	Example of MRI brain scan from a musician . . . . .	41
4.1	Example of a 3-D brain MRI scan: (a) Sagittal (SAG) plane, (b) Transverse (TRA) plane and (c) Coronal (COR) plane . . . . .	43
4.2	Example of the MRICro software interface showing three different views of a MRI brain scan image . . . . .	45
4.3	Examples of 3-D MRI brain scan slices in the Sagittal (SAG) plane . . . . .	46
4.4	Examples of 3-D MRI brain scan slices in the Sagittal (SAG) plane after applying thresholding (threshold value = 0.30) . . . . .	46
4.5	Illustration of the Volume Growing 3-D Segmentation Algorithm in the Sagittal (SAG) plane . . . . .	48
4.6	Illustration of the Volume Growing 3-D Segmentation Algorithm in the Transverse (TRA) plane . . . . .	49

4.7	Pre-defined target areas (Bounding Boxes) required for the Bounding Box 3-D segmentation algorithm with respect to MRI brain scan slices in the Sagittal (SAG) plane . . . . .	51
4.8	Example of an extracted right ventricle generated using Meshlab . . . . .	51
4.9	Example of a MRI brain scan - COR plane . . . . .	52
4.10	Comparisons between volumes obtained using manual estimation and the Bounding Box algorithm in the Sagittal (SAG) plane . . . . .	53
4.11	Comparisons between volumes obtained using manual estimation and the Bounding Box algorithm in the Transverse (TRA) plane . . . . .	54
4.12	Comparisons between volumes obtained using manual estimation and the Volume Growing algorithm in the Sagittal (SAG) plane . . . . .	55
4.13	Comparisons between volumes obtained using manual estimation and the Volume Growing algorithm in the Transverse (TRA) plane . . . . .	56
4.14	Bland-Altman plot comparing volumes obtained using manual estimation and the Bounding Box 3-D segmentation algorithm in the Sagittal (SAG) plane . . . . .	57
4.15	Bland-Altman plot comparing volumes obtained using manual estimation and the Bounding Box 3-D segmentation algorithm in the Transverse (TRA) plane . . . . .	58
4.16	Bland-Altman plot comparing volumes obtained using manual estimation and the Volume Growing 3-D segmentation algorithm in the Sagittal (SAG) plane . . . . .	59
4.17	Bland-Altman plot comparing volumes obtained using manual estimation and the Volume Growing 3-D segmentation algorithm in the Transverse (TRA) plane . . . . .	60
5.1	Average classification accuracy of Statistical metrics based classification for the Epilepsy dataset . . . . .	67
5.2	Average classification accuracy of Statistical metrics based classification for the Musician dataset . . . . .	68
5.3	Average classification accuracy of Statistical metrics based classification using SVM . . . . .	69
5.4	Average run time complexity of Statistical metrics based image classification . . . . .	70
5.5	The total run time using Statistical metrics based classification (representation generation time plus classification time) . . . . .	71
6.1	Representation of a ventricle (VOI) as a point series using the Disc-based technique . . . . .	75

6.2	An example of a point series curve generated using the Disc-based technique . . . . .	76
6.3	Representation of a ventricle (VOI) as a point series using the Spoke-based technique . . . . .	76
6.4	An example of a point series curve generated using the Spoke-based technique . . . . .	77
6.5	Example of a curve $Z$ . . . . .	79
6.6	Global warping path constraints . . . . .	82
6.7	Curve generation time for the Epilepsy dataset . . . . .	85
6.8	Classification time for the Epilepsy dataset . . . . .	85
6.9	Curve generation time for the Musician dataset . . . . .	86
6.10	Classification time for the Musician dataset . . . . .	86
6.11	The relation between classification accuracy and spoke spacing for the Epilepsy dataset . . . . .	87
6.12	The relation between classification accuracy and spoke spacing for the Epilepsy dataset with meta attributes . . . . .	87
6.13	The relation between classification accuracy and spoke spacing for the Musician dataset . . . . .	88
6.14	The relation between classification accuracy and spoke spacing for the Musician dataset with meta attributes . . . . .	88
6.15	Classification time for the Epilepsy dataset with meta attributes . . . . .	91
6.16	Classification time for the Musician dataset with meta attributes . . . . .	91
7.1	Schematic illustrating the graph/tree based approach . . . . .	93
7.2	Illustration of the Oct-tree hierarchical decomposition process given a VOI surrounded by a MBB . . . . .	94
7.3	Example dataset for illustrating the process of calculating weights using the ATW scheme . . . . .	100
7.4	Run time complexity of Oct-tree generation and frequent subgraph mining process using gSpan . . . . .	105
7.5	Run time complexity of image classification for the Epilepsy dataset using SVM classification . . . . .	105
7.6	Run time complexity of image classification for the Epilepsy dataset using Decision Tree classification . . . . .	106
7.7	Run time complexity of image classification for the Musician dataset using SVM classification . . . . .	107
7.8	Run time complexity of image classification for the Musician dataset using Decision Tree classification . . . . .	107

8.1	Run time complexity for the classification process for the Epilepsy dataset (with and without augmentation) . . . . .	112
8.2	Run time complexity for the classification process for the Musician dataset (with and without augmentation) . . . . .	113
8.3	Critical difference diagram for the proposed image classification approaches	115

# List of Tables

5.1	Classification results of Statistical metrics based classification for the Epilepsy dataset . . . . .	66
5.2	Classification results of Statistical metrics based classification for the Musician dataset . . . . .	66
5.3	Classification results of Statistical metrics based classification for Epilepsy and Musician dataset using SVM . . . . .	67
5.4	Total run time (seconds) for classification (model generation and testing)	69
5.5	The total run time using Statistical metrics based classification (representation generation time plus classification time) . . . . .	70
5.6	Classification results using Statistical metrics based classification, single attribute feature vectors and SVM . . . . .	71
6.1	Example of the accumulator matrix generated from the curve $Z$ in Figure 6.5 . . . . .	79
6.2	Example of Matrix $A$ considering distance between curve $X$ and $Y$ . . .	82
6.3	Classification results for the Epilepsy dataset using Feature space classification . . . . .	83
6.4	Classification results for the Musician dataset using Feature space classification . . . . .	84
6.5	Classification results for the Epilepsy dataset using KNN classification .	84
6.6	Classification results for the Musician dataset using KNN classification .	84
6.7	Classification results for the Epilepsy dataset with meta attributes using the Feature space classification . . . . .	89
6.8	Classification results for the Musician dataset with meta attributes using the Feature space classification . . . . .	90
6.9	Classification results for the Epilepsy dataset with meta attributes using the KNN classification . . . . .	90
6.10	Classification results for the Musician dataset with meta attributes using the KNN classification . . . . .	90
7.1	Notation used throughout this chapter . . . . .	97

7.2	Classification results for the Epilepsy dataset using SVM and Decision Tree . . . . .	102
7.3	Classification results for the Musician dataset using SVM and Decision Tree . . . . .	103
7.4	Classification results for the Epilepsy dataset augmentation and using SVM and Decision Tree . . . . .	104
7.5	Classification results for the Musician dataset augmentation and using SVM and Decision Tree . . . . .	104
8.1	Best classification effectiveness results for Epilepsy dataset (without augmentation) . . . . .	110
8.2	Best classification effectiveness results for Musician dataset (without augmentation) . . . . .	110
8.3	Best classification effectiveness results for Epilepsy dataset (with augmentation) . . . . .	110
8.4	Best classification effectiveness results for Musician dataset (with augmentation) . . . . .	111
8.5	Area under the receiver operating characteristic Curve (AUC) results . .	114
8.6	Comparison of AUC between the work reported in this thesis (3-D) and previous work (2-D) . . . . .	115

# Chapter 1

## Introduction

Data mining has been a popular area of study within the domain of Knowledge Discovery in Data (KDD), which in turn is the non-trivial process of identifying valid, useful and understandable information in data [14]. Data mining is an essential element within the overall KDD process that is concerned with the actual discovery of hidden information within data. An important activity within the remit of data mining is classification, the construction of models to categorise previously unseen data. The construction of classification models can be conducted in either a supervised or unsupervised manner. We say that the model, called a classifier, is learnt; hence we have supervised and un-supervised learning. The distinction is that in the case of supervised learning the classifier is built using pre-labelled training data while in the case of un-supervised learning no such training data exists. (There is also the concept of semi-supervised learning, but this is outside of the scope of this discussion.)

Data mining can be applied in the context of different forms of data such as document collections, graphs, videos and so on. With respect to the work presented in this thesis, the data of interest is image data, particularly 3-D image data. Data mining when applied to image data is commonly referred to as image mining. Image mining involves a number of challenges of which the most significant relates to the representation of the image data [63]; the data mining techniques themselves are well understood. The representation must be: (i) succinct (efficient), (ii) serve to capture sufficient detail (accurate) and (iii) allow for the application of some form of data mining technique (practical). The broad domain of image mining can be divided into *whole image* mining and *Region Of Interest* (ROI) image mining where the distinction is that the latter is directed at some specific sub-image that exists across an image collection; the work presented in this thesis is directed at the latter.

More specifically the work presented in this thesis is directed at 3-D ROI image mining, in other words, Volume Of Interest (VOI) image mining. The particular branch of data mining of interest is supervised classification, hence VOI image classification. To act as a focus for the research presented in this thesis the work is directed at 3-D

Magnetic Resonance Imaging (MRI) brain scan VOI classification. As such the work can be said to build on established work on 2-D MRI brain scan ROI classification [40, 140, 98]. Another challenge associated with the mining of 3-D image data, when compared to 2-D image data, is that there is a significant (exponential) increase in the amount of data to be considered.

The broad aim of the research presented in this thesis was to investigate feature-based representations appropriate for use with 3-D image classification systems so as to support the concept of 3-D volumetric data mining. The philosophical underpinning supporting the proposed work is founded on the observation that to produce an accurate VOI classifier a representation that can be interpreted by human observers is not required; what is required is a representation that results in effective classification.

In supervised image classification a collection of pre-labelled images is used as input in order to train a classifier (model) which can then be applied to unlabelled images. Image representation, as already noted, is a critical precursor to classifier generation where the images of interest are required to be transformed into a format whereby established classification techniques can be applied. In the case of supervised image classification, when applied to VOI, it is necessary to first identify the VOI (ROI) prior to considering any form of representation. Some form of segmentation is therefore required. In this thesis the term VOI Based Image Classification (VOIBIC) is used to describe the supervised learning process of building and applying classifiers for labelling VOI within 3-D data collections.

The rest of this introductory chapter is organised as follows. The motivation for the research is presented in Section 1.1. The research objectives, including the fundamental research question to be addressed and the associated research issues, are given in Section 1.2. In Section 1.3 the adopted research methodology used to provide an answer to the research question and associated research issues is presented. Next, the contributions of the research are presented in Section 1.4. Finally, the organisation of the rest of this thesis, and some concluding remarks concerning this chapter, are presented in Sections 1.5 and 1.6 respectively.

## 1.1 Motivation

From the above, the research described in this thesis is directed at 3-D supervised image classification or VOIBIC. There are many applications where 3-D image classification can be utilised; however, with respect to the research presented in this thesis, the focus is on medical applications. Automated (or semi-automated) medical prediction is a challenging real world problem [77, 10, 113]. The effective and efficient automated prediction of medical conditions is clearly of significant benefit especially with respect to speed of diagnosis and resource management, even if used as a “first screening”. Medical imaging technologies, such as Magnetic Resonance Imagery (MRI), Magnetic Resonance

Spectroscopy (MRS) or Computed Tomography (CT), have developed rapidly over the last two decades [33]. As a consequence the quantity of 3-D medical image data that is available for diagnosis has also grown significantly. As a result, sophisticated tools to help and assist in the analysis of medical image data have become a requirement. One mechanism whereby the 3-D analysis task can be supported, and that of interest with respect to this thesis, is to use classification techniques in order to classify 3-D medical image data, in other words VOIBIC. This is then the main motivation for the work described in this thesis. The principle challenge of VOIBIC is the identification of discriminative features within the 3-D data of interest.

The application focus for the work presented in this thesis is the analysis (supervised classification) of VOIs within 3-D MRI brain scan volumes. 3-D MRI brain scan data consists of a sequence of 2-D “slices” in three planes: Sagittal (SAG), Coronal (COR) and Transverse (TRA). Collectively this set of slices is referred to as a *volume*. There are a number of features in MRI brain scan volumes that are of interest. One example of previous work that has been directed at 2-D ROI analysis of MRI brain scan data is the application of classification techniques to the corpus callosum [39, 40, 37, 47, 133], the central part of the brain that connects the two halves of the brain together. The work presented in this thesis is directed at the lateral ventricles of the brain. This 3-D feature was chosen because a number of studies had suggested that there are correlations between the structure and size of the brain ventricles and some medical conditions such as epilepsy [79, 28, 18, 69, 115].

In summary, from a technical perspective, the work described is motivated by the need for techniques that can represent and classify 3-D medical images according to the structure of the features of the VOIs that occur across such medical image sets. From an application perspective the work is motivated by a desire to analyse the left and right ventricle VOI found in 3-D MRI brain scan data because it is suggested that the shape and size of the ventricles has an effect on medical conditions such as epilepsy.

## 1.2 Thesis Objectives

As a consequence of the motivation given in the previous section, this thesis is focused on an investigation of techniques to facilitate the analysis of 3-D medical images in terms of particular VOI that occur across an entire image collection. More specifically, the main objective of the work described in this thesis is directed at an investigation of automated techniques for classifying 3-D MR brain scan images (or volumes) in terms of the left and right ventricle VOI that feature in such volumes. The research questions to be addressed in this thesis are thus:

1. **What is the most appropriate mechanism for identifying VOI, and in particular the left and right ventricle VOI?:** Clearly some form of 3-D

image segmentation is required, but at the outset of the research the nature of this segmentation was unclear.

2. **Assuming that we can establish a process for identifying the VOI, how do we ascertain the quality of this VOI?:** In the case of the left and right ventricle VOI there is no “gold standard”, the mechanism whereby a VOI segmentation can be verified remains a research challenge.
3. **Once the VOI have been identified what would be the nature of the VOI representations to be used so as to support classification?:** A variety of different mechanisms can be envisioned whereby the segmented VOI can be represented. At the commencement of the research it was unclear what the most appropriate representation might be. As already noted above, the primary criteria for any proposed representation is classification effectiveness (accuracy), however there is also a requirement that the representation is efficient in terms of processing time.
4. **What are the most appropriate parameter settings for the considered representations?:** Any representation generation mechanism will entail usage of a set of parameter settings, what the most appropriate settings are was a subject for further investigation within the context of this thesis.
5. **Given a specific representation what is the most appropriate classification model to be used with that representation?:** A great number of methods have been proposed for conducting classification with respect to a similar variety of data representations. At commencement of the research, and in the absence of any identified representation mechanism, it was unclear as to which classification model would be the most appropriate. This was therefore another element of the research presented in this thesis.
6. **What is the overall most appropriate VOI classification process in the context of the ventricle application?:** It can be anticipated that different classification techniques will be best suited to different representations. However, can a “best performing” approach be identified? And if so is this “best” performance statistically significant?
7. **Given a VOI representation is there any benefit to be gained by augmenting the data?:** With respect to many application domains it can be anticipated that some additional meta data will be available. In the case of the ventricle data the age and gender of each subject was available. Whether augmenting the VOI representations with this extra meta data would provide benefits in terms of classification effectiveness was unclear at the start of the research presented in this thesis.

### 1.3 Research Methodology

Given the above research objectives and issues, the adopted research methodology was to first consider how best to identify the VOI in the context of the MRI ventricle classification application domain. The idea was to consider a number of segmentation techniques and realise the most promising. The start point for this was a review of current 3-D segmentation techniques. Most of these current techniques are extensions of existing 2-D techniques where there has been much more reported research work. It was anticipated that some form of “region growing” technique would be the most appropriate.

A significant challenge when considering the appropriateness of 3-D segmentation techniques is the quality of the resulting segmentation. As noted above, in the case of the left and right ventricle application, the main application motivation for the research presented in this thesis, there is no “ground truth” (“gold standard”). The only way of comparing the validity of any proposed segmentation algorithm was to compare the results with those produced using “hand-crafted segmentation”. Some form of statistical evaluation would thus be required.

Once the VOI had been identified the next stage was to consider the representation of the VOI. There are a variety of image representation techniques that can be adopted, the majority proposed in the context of 2-D image data. A preliminary review of the existing work on 2-D and 3-D image approaches suggested three broad categories of technique: (i) Statistical Metrics, (ii) Point series and (iii) Oct-tree:

1. **Statistical Metrics:** This was considered to be the most straightforward mechanism for representing VOIs. The idea was to represent VOIs in terms of a set of statistical metrics features such as: length, height, width, perimeter length, area, convex area, solidity (area/convex area), eccentricity, or equivalent diameter. The result could then be used to define a standard feature space from which feature vectors can be extracted in the same manner as in the case of tabular data. As a consequence, standard classification models can be applied.
2. **Point series:** There has been extensive work, mostly in connection with 2-D ROI, founded on the concept of representing the outline of a ROI simply as an ordered sequence of points referenced to a 2-D coordinate system. The idea was thus to investigate the usage of such point series in terms of 3-D VOI. A challenge here was how best to impose an ordering on the 3-D referenced points. A second anticipated challenge was how to apply existing classification techniques to a 3-D point series representation. One avenue for investigation was mechanisms for converting point series into a vector space format compatible with many classification models. Alternatively some form of K-Nearest Neighbour (KNN) classification [16] could

be applied coupled with a time series similarity measures such as warping path distance [11].

3. **Oct-tree:** Tree based representations have been widely used in connection with 2-D image representations. They tend to be coupled with a particular approach to image decomposition. In the context of 2-D image data the most common approach is to use a quad-tree representation, which in turn is directly related to a quad-tree decomposition. The quad-tree thus seemed like a good start point for any investigation of tree based representations of 3-D VOI, although in 3-D an “Oct-tree” would clearly be required. A challenge would be the classification of tree (graph) represented VOIs. Some form of graph mining suggested itself. Alternatively some mechanism whereby the trees could be translated into a vector space format also seemed like a fruitful avenue for research. One idea was to use some form of Frequent Subgraph Mining (FSM) [136] in order to identify frequently occurring subgraphs which could then be used to define a feature space.

To evaluate the effectiveness of any proposed representation, of the form identified above, two 3-D MRI brain scan datasets were used, (i) Epilepsy and (ii) Musician. The Epilepsy dataset consisted of 210 MRI brain scans, one half from epilepsy patients and the other half from healthy people. The Musician dataset in turn consisted of 160 MRI brain scans of which one half was from musicians and the other half from non-musicians. The evaluation was conducted using the standard metrics frequently used in the context of classification. In the context of effectiveness it was decided to adopt the following: (i) accuracy, (ii) sensitivity and (iii) specificity. With respect to efficiency it was decided that runtime was the appropriate measure. To determine whether the outcomes were statistically significant or not it was decided to apply the Friedman Statistical Significance Test [45] and support this with the use of critical difference diagrams (as suggested in [32]) so that statistically relevant distinctions (if any) between the techniques proposed could be highlighted.

## 1.4 Thesis Contributions

The main contributions of the research presented in this thesis can be summarised as follows:

1. Two segmentation algorithms, the Volume Growing and Bounding Box algorithms, to extract a particular VOI across an image dataset. An image thresholding technique was also applied to help partition the images according to foreground and background.
2. An approach to 3-D MRI classification using a representation based on statistical metrics describing the geometry of a detected VOI. More specifically using geo-

metrical features such as: width, length, height, perimeter (with respect to the three primary axes), volume and volume extent, of the VOI.

3. An approach to 3-D MRI classification based on a point series representation coupled with Hough signature extraction. The latter used to produce signatures from the generated point series curves, which could then be used as the input to a classification system.
4. An approach to 3-D MRI classification based on a point series representation coupled with the well known K-Nearest Neighbour (KNN) algorithm and Dynamic Time Warping (DTW) to generate a “Warping Distance” as the similarity measure.
5. An approach to 3-D MR classification based on an Oct-tree representation of an identified VOI. The Oct-tree hierarchical representation was coupled with a Frequent Subgraph Mining (FSM) algorithm to identify frequently occurring subtrees (within the Oct-tree representation) which were then used to define a feature space from which feature vectors could be extracted to which an “off the shelf” classification model could be applied.
6. An effective framework for classifying the nature of the left and right ventricles in human MRI brain scans in order to support various kinds of diagnosis and analysis.

## 1.5 Thesis Organisation

The organisation of the rest of the thesis is as follows. In Chapter 2, an extensive literature review concerning medical image classification, including image segmentation, is presented. Chapter 3 explains the nature of the MRI brain scan datasets and the application domain used as the focus for the work described in this thesis. Chapter 4 explains the adopted image pre-processing and proposed VOI segmentation techniques, and includes an evaluation of these techniques. In Chapters 5 to 7 the three proposed volumetric image classification techniques are described in detail as follows: (i) Chapter 5 describes image classification using the proposed Statistical feature representation; (ii) Chapter 6 describes the proposed Point-series image classification technique, including usage of “Dynamic Time Warping” for similarity computation and the “Hough” signature extraction mechanism; whilst (iii) Chapter 7 presents the Tree based image classification technique together with the “Oct-tree” decomposition model and the frequent subgraph mining technique used to define the desired feature space. Chapter 8 then presents a comprehensive evaluation of the proposed techniques including a statistical significance analysis of the results obtained. Finally, Chapter 9 concludes the thesis with: (i) a summary of the contents, (ii) a review of the main technical and

application based research findings in the context of the research questions listed in Section 1.2 above and (iii) some prospective future research directions.

## **1.6 Conclusion**

This chapter has provided the necessary context, motivation and background for the research presented in this thesis. The research objectives, methodology and contributions have all been introduced. In the following chapter a detailed literature review of the related background and existing previous work, with respect to this thesis, is presented.

## Chapter 2

# Background and Related Work

This chapter presents a review of the background and related work with respect to the research described in this thesis. The related work is mainly founded on three areas of research study as shown in Figure 2.1: (i) Magnetic Resonance Imaging (MRI), (ii) image representation and (iii) data mining. Each of these is considered in this chapter. The chapter commences with a review of the application domain in Section 2.1, namely: (i) the nature of MRI brain scans and (ii) the medical conditions whose diagnosis is associated with MRI brain scans. In the context of image representation the concepts of image segmentation and image representation, particularly in 3-D, are reviewed in Section 2.2 and Section 2.3 respectively. Section 2.4 then considers data mining and provides a review of the standard image classification mechanisms adopted with respect to this thesis. This is followed in Section 2.5 with a discussion of the literature concerning MRI brain scan classification to date. A review of the evaluation techniques used later in this thesis is presented in Section 2.6. Finally the chapter is concluded with a summary in Section 2.7.

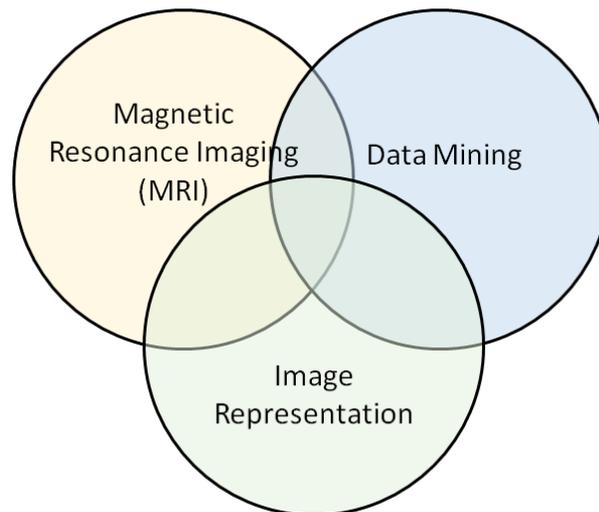


Figure 2.1: The main research areas of study related to the work present in this thesis: Magnetic Resonance Imaging (MRI), image representation and data mining

## 2.1 Application Domain

As noted in the previous chapter, the focus of the work presented in this thesis is the classification of 3-D MRI brain scans. This section reviews this application domain. The section consists of two sub-sections: (i) Magnetic Resonance Imaging (MRI) and (ii) medical conditions relating to MRI brain scan data.

### 2.1.1 Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging (MRI) came into prominence in the 1970s. It is a medical imaging technique used to investigate the anatomy and physiology of the (animal and human) body, allowing the user to “see” inside without the need for dissection [35]. MRI is similar to Computerised Topography (CT) where a scanned image is generated from many cross-sectional images. MRI uses strong magnetic fields and radio waves to produce high quality and detailed computerised images of the inside of some object. MRI is based on the principle of Nuclear Magnetic Resonance (NMR), a spectroscopic technique used by scientists to obtain microscopic chemical and physical information about molecules. In 1977, Raymond Damadian conducted the first MRI examination of a human subject [31]. MRI started out as a tomographic imaging technique that produced an image of the NMR signal in a thin slice through a body. Since then, MRI has advanced beyond a tomographic imaging technique to a 3-D imaging technique. MRI nowadays is commonly used to examine the spine, joints, abdomen, pelvis and brain of subjects. MRI can produce a very detailed image of the scanned object that can be viewed as a 3-D image.

An MRI body scanner is a short cylinder that is open at both ends. The patient lies horizontally on a motorised bed that can be moved into the scanner cylinder. It is important that the part of the body to be scanned remains still during the scan in order to produce a quality MR image. For the MRI brain scan, a “receiver frame is placed over the head of the scanned person. This frame contains “receivers” that read the magnetic resonance signals sent out by the head of the scanned person during the scan. An example of an MRI scanner is shown in Figure 2.2 and an example of a MRI brain scanner, with a receiver frame, is shown in Figure 2.3.<sup>1</sup> The process takes approximately 20 minutes. Examples of the MR brain scan images are given in Chapters 3 and 4.

---

<sup>1</sup>Photos from Center for Advanced Brain Imaging, Georgia State University.



Figure 2.2: Example of a MRI scanner



Figure 2.3: Example of a MRI brain scanner with a “receiver frame”

As noted in the previous chapter, the focus of the work described in this thesis is the classification of 3-D MRI brain scans according to a particular feature (VOI) within these scans, namely lateral ventricles. The nature of the MRI brain scan datasets used for evaluation purposes later in this thesis is discussed in Chapter 3. The medical conditions relating to MRI brain scans, epilepsy in particular, are described in the following sub-section.

### 2.1.2 Medical Conditions Relating to MRI Brain Scan

As already noted, the focus of the work described in this thesis is directed at the classification of 3-D MRI brain scans with respect to the size and shape of the lateral ventricles. The ventricles are fluid-filled open spaces at the centre of the brain [79]; there are in fact four ventricles in a human brain, but only the lateral ventricles were considered with respect to this thesis. This was because the lateral ventricles can be easily distinguished within 3-D MRI brain scan.

The size and shape of organs in human brains has been shown to be correlated to certain medical conditions or diseases (such as epilepsy, schizophrenia and multiple sclerosis) and various lateralised behaviour in people (such as handedness). It is also conjectured that the size and shape of organs in the human brain reflect certain human characteristics (such as mathematical or music ability). Within neuroimaging research considerable effort has been directed at quantifying parameters such as length, surface area and volume of structures in human brains, and investigating differences in these parameters between sample groups.

Several studies indicate that the size of the ventricular system (including lateral ventricles, third ventricle and fourth ventricle) in humans is correlated to some medical conditions. For example, in 1917, Thom was the first to note ventricular enlargement in patients with epilepsy, a medical condition whereby nerve cell activity in the brain is disturbed. Epilepsy causes abnormal behaviour accompanied by symptoms such as loss of consciousness or convulsions. Thom reported lateral ventricle dilation following postmortem examinations of patients with idiopathic epilepsy [120]. Epilepsy is significant with respect to the work presented in this thesis because one of the evaluation data sets used focuses on Epilepsy.

Over the decades the ventricular system has also been associated with other disorders, such as multiple sclerosis, schizophrenia, Alzheimers Disease (AD) and Parkinsons Disease (PD) [69]. Ventriculomegaly is commonly observed in most neurodegenerative disorders and results from passive enlargement of the lateral, third and fourth ventricles following brain parenchymal shrinkage. Significant ventricular enlargement has been associated with AD [7]. Similarly, an explorative study of PD suggests that ventricular enlargement is associated with early cognitive impairment. Previous work has also shown that ventricular enlargement (of the lateral and third ventricles) was associated with neuropsychological functions in advanced non-demented PD patients [29]. Furthermore, the presence of ventricular enlargement in both epilepsy and schizophrenia has indicated a common neurodevelopmental mechanism that predisposes to epileptogenesis and schizophrenia [18]. In a study of severe myoclonic epilepsy in infancy, 6 out of the 13 patients investigated exhibited moderate ventricular enlargement [115]. Age-accelerated changes in epilepsy participants (in comparison with healthy people) have been seen in the lateral ventricles, whereas largely comparable patterns of age-related

changes were seen across other regions of interest in the brain [28].

Although the work described in this thesis is directed at MRI brain scan classification focusing on the lateral ventricles, there are other parts in MRI brain scans to which the techniques could be applied, such as the corpus callosum or hippocampus.

## 2.2 Review of Image Segmentation

In computer vision, image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels) in order to simplify or change the representation of the image that is more meaningful or easier to analyse [109]. With respect to the work presented in this thesis we are interested in particular segments. Prior to the application of any segmentation process some form of image pre-processing, such as image registration and noise removal, is required. Image registration is the process of insuring that a set of images conforms to a single coordinate system. This is important so that a collection of images can be effectively compared. Image registration can be defined as the problem of identifying a set of geometric transformations which map the coordinate system of an image set to that of the others. With respect to the work in this thesis, image registration was conducted to ensure that all MRI brain scans in a given collection conformed to the same coordinate axes before the segmentation process is commenced. Image noise reduction is the process of removing “noise” from images; unwanted pixels that adds spurious and extraneous information to a given image.

For 3-D image segmentation, there are many techniques that can be used in the context of brain MRI scan segmentation such as: Thresholding, Region Growing and Deformable Models and Level Sets. Each is briefly discussed in the following subsections.

### 2.2.1 Image Segmentation Using Thresholding

Thresholding is an image segmentation method based on the assumption that the foreground of an image can be categorised by its brightness. Thresholding is arguably the most widely used of all segmentation methods [132]. There are three major thresholding segmentation techniques: (i) global thresholding, (ii) local thresholding and (iii) Hysteresis thresholding, described as follows:

#### Global Thresholding

Global thresholding is the simplest thresholding segmentation technique. Given a function  $f(x)$  that returns the brightness level of a given image voxel  $x$ . a threshold value  $\theta$  can be defined such that  $\min_x(f(x)) \leq \theta \leq \max_x(f(x))$ . We can then define a second function  $g(x)$  that returns 1 if the given voxel  $x$  is to be deemed part of the foreground, and 0 otherwise, as follows:

$$g(x) = \begin{cases} 1 & \text{if } f(x) \geq \theta \\ 0 & \text{else} \end{cases} \quad (2.1)$$

While Equation 2.1 is a complete description of a binary algorithm, it contains no indication how to select the value of  $\theta$ . The natural question is whether there exists an optimal threshold value. There are different solutions to this threshold selection problem, each solution is based on a different model of assumptions. However even if  $\theta$  was optimally selected using a model which is suitable for the present image data, global thresholding can give poor results whenever the influences of noise are large compared to the image content, or when object and background grey value intensities are not constant throughout the volume. Some solutions for global threshold selections are as follows:

1. **Choosing thresholds using prior knowledge:** In many situations an image objected to be segmented has known physical properties. The optimal threshold value  $\theta$  can then be defined accordingly.
2. **Otsu’s method:** Proposed by Otsu [93], the idea is to choose a threshold value  $\theta$  by analysing the distribution of grey values of an image. Assuming that the grey value histogram of a given image contains two separate peaks, one indicating foreground voxels and the other background voxels, the minimum between these two peaks can be selected as the threshold  $\theta$ .
3. **Isodata method:** The Isodata method is described in [102], the idea is to apply a general *Isodata clustering* algorithm to the grey values of an image. The Isodata clustering algorithm is a user-defined threshold algorithm. The threshold is then somewhere between the means of the foreground and background (as in the case of Otsu’s method), but instead of searching for a global optimum as in Otsu, the search is performed locally.

### Local Thresholding

Local thresholding can solve some of the shortcomings associated with global thresholding as described above. The intensity levels of an image can vary depending on the location within a data volume. One common local thresholding approach is to calculate the mean intensity values within a window around each pixel and subtract these sliding mean values from each pixel.

$$g(x) = \begin{cases} 1 & \text{if } f(x) \geq \theta(x) \\ 0 & \text{else} \end{cases} \quad (2.2)$$

As was the case for the global thresholding approach, Equation 2.2 does not provide any clues on how the threshold values  $\theta(x)$  should be computed. Some common approaches are as follows:

1. **Niblack Thresholding:** In 1995, Trier et al [121] proposed a local thresholding technique using Niblack’s algorithm [86] to calculate local mean and standard deviations to obtain a threshold. The method implicitly assumes smooth foreground and background areas, where the grey values vary about some unknown mean, which is estimated in a window around a current coordinate point. The approach worked well in practice when the window size can be chosen to correspond to the size of objects that are present in a given image. The system is likely to fail, on the other hand, in large low-contrast areas.
2. **Mardia and Hainsorth Method:** Mardia et al. proposed an algorithm for local thresholding in 1988 [83]. Their idea was to obtain random variables  $G(x)$  for each voxel location  $x$  and define these as linear combinations of the neighbouring voxels  $x'$ . The selection of a local threshold  $\theta(x)$  was then conducted using the “Mardia-Hainsworth” algorithm described in [83]. The disadvantage of this method is that the cost of computation is significantly high comparing to the other segmentation methods.
3. **Indicator Kriging:** Thresholding by Indicator Kriging was described by Oh et al. in 1999 [91]. Kriging is an interpolation method that is commonly used in geostatistics. It is similar to the Mardia and Hainsorth method in that it estimates the value at voxel  $x$  using a linear combination of its neighbours. This method is a modification of the Mardia and Hainsorth method described above; the main difference is that the Indicator Kriging method uses covariance estimation to calculate the local threshold.

### Hysteresis Thresholding

Another common problem in image segmentation is that the segments of interest may be defined by their intensities, but that there also exist other structures (noise) with high values. Global thresholding would either underestimate the size of the true segments (because  $\theta$  was too high) or would include noise in the foreground (because  $\theta$  was too low). One way of dealing with such situations where the voxel value distributions of foreground and background voxels overlap is Hysteresis thresholding, also known as “double thresholding” proposed in [17]. Hysteresis thresholding uses two thresholds,  $\theta_1 > \theta_2$ , and starts from voxels  $x$  with  $f(x) \geq \theta_1$ . Then all voxels  $x'$ , which are neighbours of an already identified foreground voxel and which fulfil the condition  $f(x') \geq \theta_2$ , are iteratively assigned to the foreground. This procedure ensures segmentation of connected regions, since a number of “certain” foreground elements are selected while its

neighbours may have a lower value. At the same time, noisy background voxels are suppressed by the higher threshold  $\theta_1$ . Note that this algorithm has some similarity to the region growing algorithms that are introduced in the following sub-section.

### **2.2.2 Image Segmentation Using Region Growing**

The concept of all region growing for segmentation is that voxels belonging to one object are connected and similar. The common procedure is to compare one voxel with its neighbours. If a similarity criterion is satisfied, a given voxel can be set to belong to the cluster associated with one or more of its neighbours. The selection of the similarity criterion is significant and the results are influenced by noise in all instances. Some well-known region growing segmentation methods are listed below.

#### **Growing by Value**

Growing by Value is one of the most commonly used region growing segmentation methods. The method is founded on the observation that an object's grey values are usually within some range around a mean value. Thus, while growing a region, its current mean and standard deviation are computed and a new voxel is added to the region if its value is within a range around the regions mean. It can produce reasonable segmentation results where objects are connected and can be characterised by their grey values.

#### **Adaptive Region Growing**

The Adaptive Region Growing method was proposed in [84] for segmentation of the human cortex. Their idea was to adapt the "decision function" according to the regions size. Initially, for a region containing very few voxels, voxels are added as long as a homogeneity (grey value variance) threshold around the region is not exceeded. Then, when a certain number of voxels has been added to the region, it is assumed that the grey level statistics of this region have approached the objects true distribution.

#### **Adams Seeded Region Growing**

The Adams Seeded Region Growing method [4] takes a set of seeds (voxels) as input along with the image. The seeds mark each of the objects to be segmented. The regions are iteratively grown by comparison of all unallocated neighbouring pixels to the regions. The difference between a pixel's intensity value and the region's mean,  $\delta$ , is used as a measure of similarity. The pixel with the smallest difference intensity is assigned to the respective region. This process continues until all pixels are assigned to a region. Because Seeded Region Growing requires seeds as additional input, the segmentation results are dependent on the choice of seeds, and noise in the image can cause the seeds to be poorly placed.

## Non-Connected Region Growing

A region growing algorithm that can segment non-connected regions was proposed by Revol et al. [100]. The unique feature of this algorithm is that voxels may not only be added, but also removed from a region. To achieve this, the so-called “ $k$  – contraction” is used.  $k$  – contraction removes  $k$  voxels from a region, starting from the voxel with the lowest grey value, in increasing order. Note that this procedure assumes that object voxels have larger grey values than background voxels. The procedure is repeated until a homogeneous region is produced, where a region is called homogeneous if its grey value variance is below a predefined threshold.

## Parameter-Free Region Growing

The Parameter-Free Region Growing algorithm is an extension of the Non-Connected Region Growing described above that uses an “Assessment Function” [101] or “Unseeded Region Growing” [80] in order to help start region growing automatically.

### 2.2.3 Deformable Methods and Level Set Methods

The Deformable surfaces and Level Set Methods are model-based approaches to image segmentation that have widely been applied in both 2-D and 3-D medical image processing. The main methods of Deformable and Level Sets Methods are described in the following sub-sections.

#### Deformable Models

Kass et al. proposed the use of deformable models for image segmentation in 1988 [74]. Deformable models are physically motivated, model-based, techniques for delineating region boundaries by using closed parametric curves or surfaces that deform under the influence of internal and external forces. Internal forces are computed from within the curve or surface to keep it smooth throughout the deformation. External forces are usually derived from the image to drive the curve or surface toward the desired feature of interest.

#### Level Sets

The Level Sets Method was proposed by Sethian [108]. The concept is to represent the evolving contour using a signed function whose “zero” corresponds to the actual contour. According to the motion equation of the contour, it can derive a similar flow for the implicit surface that, when applied to the zero level, will reflect the propagation of the contour. The level set method affords numerous advantages: it is implicit, is parameter-free and provides a direct way of estimating the geometric properties of the

evolving structure. It has been argued to be a very convenient framework for addressing numerous applications of computer vision and medical image analysis.

#### **2.2.4 Other Segmentation Methods**

The segmentation methods that were described so far fell into three groups: (i) Grey value based (Section 2.2.1), (ii) Region based (Section 2.2.2) or (iii) Model based (Section 2.2.3). A number of other segmentation methods are considered as follows.

##### **Image Segmentation Using Fuzzy Correctedness**

The idea behind fuzzy connectedness is to represent knowledge on the connectedness of voxels by a fuzzy relation [126]. A fuzzy relation can be interpreted as a measure of similarity between two voxels. The use of the fuzzy connectedness method lies in finding the connectedness of two voxels and deciding on whether these two voxels belong to the same object or not. Algorithms for computing the fuzzy connectedness between any two points in the image domain were proposed in [126, 90]. Within the fuzzy connectedness framework, segmentation of an image reduces to thresholding of the fuzzy connectedness values. Therefore, any two voxels with values that exceed this threshold will be labeled as belonging to one image segment.

##### **Image Segmentation Using Watershed Algorithm**

The name “Watershed” is analogous to the notion of a catchment basin of a height map [112]. The watershed algorithm uses concepts from edge detection and mathematical morphology [51, 112] to partition images into homogeneous regions [129]. The method can suffer from over-segmentation, which occurs when the image is segmented into an unnecessarily large number of regions. Thus, watershed algorithms are usually followed by a post-processing step to merge separate regions that belong to the same structure [111]. In 1993, Vincent et al. [128] proposed morphological reconstructions for preprocessing the set of starting points which was found to improve the performance of the algorithm.

##### **Image Segmentation Using Bayesian Methods**

Bayesian approaches to image processing treat all involved quantities as random variables and rely on the laws of probability to derive probabilistic models for images. In Bayesian decision theory [128], costs are assigned to either correct or wrong decision based on the probabilities of occurring events. The decision that minimises the risk is taken. The risk in Bayesian decision theory is the cost times the probability of a wrong action being taken.

### 2.2.5 Image Segmentation in MRI Brain Scan Data

There has been a significant amount of published work in the context of 3-D MRI brain scan segmentation. This is because brain image segmentation is seen as an important medical diagnostic tool. MR brain images mostly frequently include undesirable features such as: noise and inhomogeneity (lack of homogeneity). Therefore, accurate segmentation of brain images can be very challenging. This sub-section reviews some of the published work on MRI brain scan segmentation.

The most well known segmentation technique used in the context of MRI brain scan segmentation is thresholding [49]. An example where this was used can be found in Suzuki et al. [116] where it was used to automatically extract soft-tissue from MRI brain scan images. More specifically [116] used an iterative thresholding algorithm for the segmentation with an optimum threshold value decided according to a “goodness” measure. Brummer et al. [15] proposed a fully automation mechanisms for detecting brain contours from 3-D brain MRI data using a histogram-based thresholding technique (following by a morphological procedure which refined the binary threshold mask image and discriminated between desired and undesired brain structures). Long et al. [82] used dynamic thresholding coupled with Modified Spectral Segmentation to segment the lateral and third ventricles from MRI brain scans. The disadvantage of the thresholding technique, as noted in sub-section 2.2.1 above, is that it is often not self-contained and the threshold value has to be adjusted depending on brightness variation of images.

As a consequence of the disadvantages of thresholding Deformable and Level Set Models have become the most widely-used techniques in medical MRI brain segmentation [95]. Examples of where they have been used can be found in [75, 72, 60, 21]. Khotanlou et al. [75] proposed a method to segment brain tumors using a constrained deformable model while Ho et al. [60] proposed a similar approach but using a Level Set technique. Kapur et al. [72] proposed a method to segment brain tissue from 3-D MRI scans using a deformable model coupled with binary mathematical morphology and expectation/maximisation segmentation. Ciofalo et al. [21] presented a technique that combined the Level Sets method with fuzzy logic to segment non-tissue brain structures. Although the foregoing produced promising results, the limitations of the technique are that: (i) they are time consuming and (ii) sometimes difficult to implement.

Apart from time complexity another major issue associated with 3-D MRI brain scan segmentation is the evaluation problem [139]. It is very challenging to evaluate the accuracy of the segmentation, in humans in particular, because it is impossible to know the exact size or shape of brain organs without an dissection. The most practical way to evaluate the segmented object is to compare the features of the segmented object and those collected manually by a domain expert.

Note that with respect to the work presented in this thesis, global thresholding

coupled with the Bounding Box algorithm was used to segment the lateral ventricles (explained in Chapter 4). The reason for this was that the MRI brain scan images used in the datasets have very low brightness variation (see details in Chapter 3). It was therefore found to be beneficial to use the techniques as they produced good segmentation result while at the same time offering efficiency advantages compared to the techniques reviewed above.

## 2.3 Review of Image Representation

Image representation is an important issue with respect to image processing. It is currently not computationally possible (even with respect to high performance computing facilities) to present images to a classification algorithm in their entirety, especially in the case of 3-D images [42]. Consequently it is necessary to process the images in such a way that the classification process is tractable while at the same time minimising information loss. Therefore, one of the fundamental challenges of image mining is to determine how the low-level pixel representation can be translated into a form suited to further processing (classification in our case). The problem is compounded in the context of 3-D images where a voxel representation is used. A voxel is a smallest distinguishable volume element in a 3-D image, which is the 3-D equivalent of a pixel. In [114]) it was noted that the representation contained in a raw image can be efficiently and effectively processed to identify high-level spatial VOIs and their relationships. Image classification techniques normally use visual content such as: colour, texture, size and shape, to represent and classify images. Colour and texture have been explored more thoroughly than shape. It is arguable that shape is a more useful property of VOIs than colour or texture. There is considerable evidence that natural images are recognised based primarily on their shape [85]. Amongst the research community interest in using shape features of VOIs for image classification is increasing considerably. However, image classification by shape is still considered to be a more challenging task compared to image classification based on other visual features [85]. In addition, the problem of shape-based image classification becomes more complex when the extracted VOIs are corrupted by noise. The problem of 3-D shape analysis has been considered by many authors, resulting in a significant number of research publications. Examples of 3-D image representations focusing on shape-analysis are presented in the following four sub-sections.

### 2.3.1 Statistical Based Image Representation

Statistical based image representation is the simplest way of representing the geometric features of a VOI. Normally the statistical based image representation is used in 2-D image analysis but it can equally well be applied to 3-D image analysis. There are

two kinds of representation in statistical based representation: (i) first-order and (ii) second-order [119]. For first-order representation, images are described using statistical functions such as mean, variance, energy and standard deviation of the image's intensity values. For the second-order representations, the relationship between the intensity value of each pixel with respect to those of its neighbours is taken into consideration [119]. In other words, relative location information is used. An example of a second-order representation is where the concept of a co-occurrence matrix [50, 52] is used to enumerate the number of times two intensity values appear in an image within a certain distance and a direction of each other. A Voxel Co-occurrence Matrix (VCM) is used in the same manner as a pixel co-occurrence matrix but with respect to 3-D images [50].

Examples of statistical metrics that could be used in 3-D image analysis are:

1. **Width:** The width of the object.
2. **Length:** The length of the object.
3. **Depth:** The depth of the object.
4. **Axis width:** The axis width of the object.
5. **Axis length:** The axis length of the object.
6. **Axis depth:** The axis depth of the object.
7. **Volume:** The volume of the object, in terms of the number of voxels of the object.
8. **Volume extent:** The volume divided by the volume of the bounding cube.
9. **The minimum perimeter length:** The minimum length of the perimeter on  $xy$ ,  $xz$  and  $yz$  axes.
10. **The maximum perimeter length:** The maximum length of the perimeter on  $xy$ ,  $xz$  and  $yz$  axes.
11. **Volume circularity:** The volume divided by the volume of the bounding circle.

The statistical metrics above are able to reflect the shape and size of the VOIs and can be used to generate feature vectors to which standard classifiers can be applied. The advantages of the usage of the statistical metrics are: (i) its ease of use, (ii) ease of implement and (iii) speed. However, the limitation is that it may not be able to reflect some complex detail of the shape of a given VOI. Thus, some important information concerning the VOI may be loss.

### **2.3.2 Discrete Wavelet Transform (DWT) Based Image Representation**

The Discrete Wavelet Transform (DWT) representation is a well known method for extracting frequency space information from non-stationary signals [58]. It is an effective tool for feature extraction, because it allows analysis of images at various levels of resolution. This technique requires a large amount of storage and is computationally expensive [73]. In order to reduce the feature vector dimension and increase the discriminative power, the Principal Component Analysis (PCA) [71] has been used.

### **2.3.3 Histogram Based Image Representation**

Using histogram-based representation method, there are a number of techniques that can be adapted. For example: (i) Simple Histograms, (ii) Histograms of Oriented Gradients (HOGs), (iii) Histograms of Local Binary Patterns (LBPs) and (iv) Histograms of Local Phase Quantisation (LPQ). Each is explained in more detail in the following sub-sections as follows.

#### **Simple Histograms**

For simple histograms, the x-axis represents the values for some image features and the y-axis represents a count of the number of times that each feature value occurs. The attribute-values are often grouped into sub-ranges referred to as “bins”. The simplest form of histogram image representation is where the x-axis represents intensity values. The histogram thus represents the number of times each intensity value, or group of intensity values, appears. The disadvantages of such simple histograms are: (i) significant information is lost, such as spatial information, because only the frequency of the intensity values are considered; and (ii) invariant problems, especially when two images have similar content but with different resolutions (in which case different histograms will be produced).

#### **Histograms of Oriented Gradients (HOGs)**

A more advanced histogram-based method is the use of Histograms of Oriented Gradients (or HOGs) [30]. Using HOGs the changes in the intensity values of the region, with respect to either the azimuth and/or zenith direction, are computed and referred to as gradients. In order to compute a gradient at each location the difference between the “left” and “right” neighbouring intensity values, in a given direction, is calculated. After this, the angles between the image gradients are computed and stored in what are called “orientation” bins. The gradient magnitudes in each orientation bin are accumulated. In the generated histogram, the x-axis represents directions and the y-axis the sum of the gradient magnitudes.

## Histograms of Local Binary Patterns (LBPs)

In order to generate LBPs, each pixel/voxel is compared to its immediate neighbours. For each comparison a ‘one’ is stored if the intensity value of the pixel/voxel is greater than the neighbour, otherwise a ‘zero’ is stored. The generated binary number from the sequence of neighbours then describes an integer value. In the generated histogram, the x-axis represents the computed integer values and the y-axis represents the frequency with which they occur. In order to generate a robust representation, it is desirable to compute rotation invariant LBPs. With respect to 2-D images it is straightforward to calculate rotation invariant LBPs because each location has only eight immediate neighbours. With respect to 3-D images the generation of 3-D rotation invariant LBPs (26 neighbours in contrast to 8 neighbours) is computationally expensive. To address this issue Zhao et al. [141] proposed the use of Three Orthogonal Plane LBPs (LBP-TOP). The LBP-TOP representation considers the calculation of LBPs only with respect to neighbouring voxels located in the  $XY$ ,  $XZ$  and  $YZ$  planes. A combination of HOG and LBP (HOG-LBP) has also been proposed and found to be a robust representation [131].

## Histograms of Local Phase Quantisation (LPQ)

The concept of histograms of Local Phase Quantisation (LPQ) was proposed in [94]. LPQ uses low frequency local Fourier transforms whereby a histogram of the quantised Fourier transform can be generated [92]. At each image location, a Short-Term Fourier Transform (STFT) is applied with respect to the immediate neighbours. After that the resulting values are quantised (a value of one is used if the value is bigger than or equal to zero, otherwise a value of zero is used). In this manner a binary encoding is computed for each image location which can then be interpreted as an integer value between 0-256 ( $b = \sum_{i=0}^8 q_i 2^{i-1}$ ). where  $q_i$  is the quantised value of a neighbouring pixel/voxel). Histograms describing the number of times that each integer value occurs are then computed, one per image.

### 2.3.4 Hough Transform

The Hough Transform is a widespread technique in image analysis. Its main idea is to transform the image to a parameter space where clusters or particular configurations identify instances of a shape under detection. Hough Transform based techniques are used for shape detection, either parametrised or generalised. Hough Transform was first introduced by Paul Hough in 1962 [62] with the aim of detecting alignments in television lines. It became later the basis of a great number of image analysis applications. Hough Transform is mainly used to detect parametric shapes in images. It was first used to detect straight lines and later extended to other parametric models such as circles or ellipses, being finally generalised to any parametric shape [8].

One major advantage offered by Hough Transform is that it is highly robust against image noise and degradation. Hough Transform is used for extracting shape signatures which can be used as a feature vector in the classification process. Hough Transform is suitable for this task because it maintains the spatial information associated with an image VOI (feature). The “classic” Hough Transform performs a mapping between the  $XY$  image space into  $\rho - \theta$  space. The transformation is  $\rho = x\cos\theta + y\sin\theta$ , where  $(x, y)$  are the coordinates of the voxel to be transformed, and  $(\rho, \theta)$  are the parameters of a corresponding line passing through the voxel. The parameter space is suitably quantised into bins and represented by means of an accumulator matrix initially set to zeros. Each pixel  $(x, y)$  can be conceptualised as a parametric curve in the new  $\rho - \theta$  space, where  $\theta$  varies from the minimum to the maximum values, giving the corresponding  $\rho$  values. The corresponding parametric positions can be stored in an accumulator matrix  $A$ , where each row corresponds to one value of  $\rho$ , and each column to one value of  $\theta$ . The cells in  $A$  are then incremented by the parametric curve. The local maxima within  $A$  then correspond to the dominant boundary lines of the VOI. Further details about the Hough Transform applied in this thesis are presented in Chapter 6.

### 2.3.5 Tree Based Image Representation

A common mechanism for representation 2-D images is to use quad tree. Tree based representations can also be used to represent 3-D image data. First of all, a basic understanding of the concept of a graph is required. A graph,  $G$ , is a structure that consists of a set of vertices (or nodes)  $V$ , and a set of edges  $E$ , and is usually denoted as  $G = (V, E)$  [13]. The term “node” will be used in the rest of this thesis to represent graph vertices because this is the terminology usually used on the context of trees (thus we have root, body, and leaf nodes, child and sibling nodes). Some for the discussion concerning graph terminology presented in Chapter 7.

Tree data structures have been widely applied in various domains, especially in image classification [19, 82, 40]. One main advantage of the tree data structure is its ability to focus on the “interesting” parts (sub-trees) of the input data, thus permitting an efficient representation of a problem and consequently improving execution times [107]. Tree data structures to represent images can be constructed in various ways, of which image decomposition is one of the most popular methods. There are many types of image decomposition, common mechanisms use data structures such as Oct-trees, Quad-trees and Scale Space representations [26]. With respect to the work described in this thesis, the Oct-tree representation was adopted to represent a 3-D volume. An Oct-tree is a tree data structure which can be used to represent a 3-D area which has been recursively subdivide into “octants” [99, 68]. Further detail is provided in Chapter 7. The following sub-section presents a review of graph mining and especially FSM. Once a graph representation has been constructed, for use with classification model

generations, we typically wish to convert the graph into a feature vector representation; One way of identifying features is to apply Frequent Subgraph Mining (FSM).

## 2.4 Review of Selected Classification Mechanisms

This section provides a review of the selected classification mechanisms used in this thesis: (i) Decision Tree (DT), (ii) Support Vector Machine (SVM), (iii) Bayesian Network (BN), (iv) Artificial Neural Network (ANN) and (iii) K-Nearest Neighbour (KNN). Each of them is described in Sub-sections 2.4.1 to 2.4.5 below.

### 2.4.1 Decision Tree

A Decision Tree (DT) is a decision support tool that uses a tree model of decisions and their possible consequences. In the context of classification, DTs are a widely used classification technique due to their simplicity, ease of understanding, explanation generation capability and interpretability. A DT is a tree structure where the root and body nodes represent alternatives while the leaf nodes represent individual classifications. More specifically each root or body node represents an attribute and has connections to child nodes, which contain potential individual attribute values or groups of values. Therefore, a DT could be called “a tree based classifier”. In a binary DT there can only be two alternatives at each root or body node; in other forms of DT there may be many alternatives. The challenge of constructing a decision tree is the selection of the attribute to be represented by each node in the hierarchy, and how to split the range of potential values that an attribute might have. Once a DT is constructed it becomes easy and straightforward to classify a new unseen data item starting from the root and finding a route through the DT until one of the leaves (classes) is reached. Generally DT construction is conducted in a top-down manner following a “greedy” search process, with no backtracking, based on a “divide and conquer” strategy where the training set is partitioned recursively into subsets according to some splitting criterion. Various splitting criteria have been proposed. Popular measures include Information Gain, Gini Index and Gain Ratio (see [55, 34]). A variety of decision tree generation algorithms have also been proposed [55, 34].

With respect to the work described in this thesis the C4.5 algorithm [97] was adopted as it has been considered to be a benchmark DT classifier throughout the data mining community. C4.5 uses Information Gain (IG) as the splitting criteria whereby the attribute with the highest information gain is selected to be used in the current node. IG is calculated using Equation 2.3:

$$IG(D, X) = S(D) - S(D, X) \tag{2.3}$$

where  $IG(D, X)$  is the information gain for dataset  $D$  with respect to attribute  $X$ .  $S(D)$  is the Entropy for dataset  $D$  and  $S(D, X)$  is the Entropy for the dataset  $D$  with respect to attribute  $X$ .  $S(D)$  is calculated using Equation 2.4

$$S(D) = \sum_{i=1}^{i=|c|} -p_i \log p_i \quad (2.4)$$

where  $p_i$  is the probability of class  $i \in c$ . Normally,  $p_i = \frac{|c_i, D|}{|D|}$  where  $|c_i, D|$  is the number of records corresponding to class  $i$  with respect to the entire dataset  $D$ . Intuitively,  $0 \leq S(D) \leq 1$ . Entropy is a measure of the homogeneity of a given dataset. If  $S(D) = 0$ , then all the records belong to the same class and therefore the outcome is certain.

$IG$  is thus a measure of the expected change in the information entropy from a prior state to a state that takes some information for a given attribute. In other words  $IG$  indicates the “importance” of a given attribute with respect to the DT construction process. In the context of Equation 2.3 the importance of an attribute is determined by identifying the entropy value of the attribute before and after splitting. The same calculation is made for the complete set of attributes and the attribute that maximises information gain selected for the DT node in question.

## 2.4.2 Support Vector Machine

A SVM is a classification system that tries to separate data of different classes by fitting a decision boundary (hyperplane), which maximise the “distance” between data representing two different classes provided the data is linearly separable [55]. A new unseen data item may then be mapped onto the same space and classified according to which side of the hyperplane it falls.

Given a dataset,  $D = \{(X_1, C_1), (X_2, C_2), \dots, (X_{|D|}, C_{|D|})\}$ , where  $X_i$  is a training data item and  $C_i$  is associated class label. A linear SVM is constructed as follows [55]:

1. **Find the optimal separating hyperplane:** The optimal separating hyperplane is the Maximum Marginal Hyperplane (MMH) that maximally separates tuples of different classes in the space. Identification of the MMH encompasses a number of steps:
  - (a) **Find separating hyperplane:** The separating hyperplane defined as:

$$W \cdot X + b = 0 \quad (2.5)$$

where  $W$  is the weight vector and  $b$  is a scalar value known as the bias,  $b$  may be thought of as an additional weight  $w_0$ . The hyperplanes that describe each side of the separating “gap” are defined as:

$$\begin{cases} H1 : W_1X_1 + W_2X_2 + \dots + W_{|D|}X_{|D|} + w_0 \geq 1 & \text{for } C_i = +1 \\ H2 : W_1X_1 + W_2X_2 + \dots + W_{|D|}X_{|D|} + w_0 \leq 1 & \text{for } C_i = -1, \end{cases} \quad (2.6)$$

where  $|D|$  is the number of tuples in the dataset  $D$ . Equation 2.6 shows that any data item that falls on or above  $H1$  belongs to class  $+1$ , and any tuple that falls on or below  $H2$  belongs to class  $-1$ . Training tuples that fall on  $H1$  or  $H2$  are known as support vectors. Equation 2.6 can be rewritten as:

$$C_i(W_1X_1 + W_2X_2 + \dots + W_{|D|}X_{|D|} + w_0) - 1 \geq 0 \forall_i \quad (2.7)$$

- (b) **Find MMH:** To find MMH the aim is to minimise  $\|W\|$ , subject to the constraint specified in Equation 2.6.  $\|W\|$  is the Euclidean norm of  $W$ . By minimising  $\|W\|$ , which is equivalent to minimising  $\frac{1}{2}\|W\|^2$ , the distance between  $H1$  and  $H2$  will be maximised. This is achieved using an optimisation algorithm with Lagrangian formulation and Karush-Kuhn-Tucker (KKT) conditions. Once identified, the MMH can be defined as a decision boundary:

$$D(X^\chi) = \sum_{i=1}^{sv} C_i \alpha_i X_i \chi + b_0 \quad (2.8)$$

where  $sv$  is the total number of support vectors,  $C_i$  is the class label for a support vector (or training tuple)  $X_i$ ,  $\chi$  is a test tuple, and  $\alpha_i$  and  $b_0$  are parameters determined by the optimisation algorithm.

2. **classify the test tuple:** To achieve this, a test tuple,  $\chi$ , is applied to Equation 2.8. If the sign of the computed results is positive,  $\chi$  is classified as  $+1$ . If the sign is negative,  $\chi$  belongs to the class  $-1$ .

The above process is used to train linear SVMs, where the training data is assumed to be linearly separable. This algorithm can be extended to learn nonlinearly separable training tuples by first transforming the nonlinear tuples into a higher dimensional space using a nonlinear kernel function. Three common nonlinear kernel functions are:

**Polynomial:**  $K(X_i, X_j) = (X_i \cdot X_j + 1)^h$

**Radial basis function:**  $K(X_i, X_j) = e^{-\gamma \|X_i - X_j\|^2}$

**Sigmoid:**  $K(X_i, X_j) = \tanh(kX_i \cdot X_j - \delta)$

Next, the constraint in Equation 2.6 is rewritten to allow errors as follows [12]:

$$C_i(W_1X_1 + W_2X_2 + \dots + W_{|D|}X_{|D|} + w_0) - 1 + \xi_i \geq 0 \forall_i \quad (2.9)$$

where  $\xi_i \geq 0$  is called the *slack variable* that allows margin errors (the hyperplanes do not separate the training tuples of different classes correctly) and misclassification. To penalise the margin error and misclassification, subject to the constraint introduced in Equation 2.9, a “soft parameter”  $C > 0$  is used to minimise  $\frac{1}{2} \|W\|^2$  as follows:

$$\min \frac{1}{2} \|W\|^2 + C \sum_{i=1}^L \xi_i \quad (2.10)$$

where  $L$  is the number of different classes.

### 2.4.3 Bayesian Network

A Bayesian network is a probabilistic graphical model. It can be used to predict the probability that a given example belongs to a particular class, in this case the network is referred to as a Bayesian classifier. Bayesian classifiers are derived from Bayes theorem, thus if  $T$  is a data tuple and  $H$  is a hypothesis that  $T$  belongs to class  $C$ , then:

$$P(H | T) = \frac{P(T | H)P(H)}{P(T)} \quad (2.11)$$

where  $P(H | T)$  is the posterior probability of  $H$  given  $T$  (i.e. it is a measure of how confident we can be that  $H$  is true given that we know  $T$  is true). Similarly,  $P(T | H)$  is the posterior probability of  $T$  given  $H$ .  $P(H)$  is the prior probability of  $H$  and  $P(T)$  is a prior probability of  $T$ . The most straight forward Bayesian classifier are founded on the Naive Bayes assumption [53]:

1. Assume a training set with  $T$  tuples and  $m$  attributes,  $A_1, A_2, \dots, A_m$ . Suppose also there are  $n$  classes,  $C_1, C_2, \dots, C_n$ . Given a tuple,  $T$ , the classifier will classify  $T$  to class  $C_i$  if and only if:

$$P(C_i | T) > P(C_j | T) \quad \text{for } 1 \leq j \leq n, j \neq i \quad (2.12)$$

where  $P(C_i | T)$  is calculated using Bayes' theorem as defined in Equation 2.11,

$$P(C_i | T) = \frac{P(T | C_i)P(C_i)}{P(T)} \quad (2.13)$$

2. Based on Equation 2.13, the probabilities  $P(A_1, A_2, A_3, \dots, A_m | C_i)$  have to be computed in order to get  $P(T | C_i)$ . Using Naive Bayes it is assumed that the attributes,  $A_1, A_2, A_3, \dots, A_m$ , are independent of one another given any class label. Thus:

$$\begin{aligned} P(T | C_i) &= P(A_1 | C_i) \times P(A_2 | C_i) \times \dots \times P(A_m | C_i) \\ &= \prod_{s=1}^m P(A_s | C_i) \end{aligned} \quad (2.14)$$

3. A tuple,  $T$ , belongs to class  $C_i$  if and only if:

$$P(T | C_i)P(C_i) > P(T | C_j)P(C_j) \quad \text{for } 1 \leq j \leq n, j \neq i \quad (2.15)$$

$P(T)$  as in Equation 2.13 is omitted from the calculation as it is constant for all classes.

#### 2.4.4 Artificial Neural Network

An Artificial Neural Network (ANN) is a mathematical model inspired by the conjectured operation of the human biological neural system. ANNs have been used with respect to a wide range of real world applications, especially in the context of environments that are continuously changing. Typically, an ANN comprises a set of layers: (i) the input layer, (ii) the hidden layer(s) and (iii) the output layer. Each layer consists of “neurons” nodes and their weighted links. The simplest structure (topology) is the input-output layer where there are no hidden layers. Note that although the complexity of an ANN structure increases as the number of hidden layer(s) increases, the effectiveness also tends to increase. The idea is to use training data to train an ANN so that the link weightings can be learnt starting with initial weights. ANNs thus fall into the supervised learning category where weights are iteratively adjusted in order to

minimise the error between the desired output and the predicted output through using a sufficient number of training examples [59]. There are various algorithms that can be used to apply the necessary weight adjustment during training. The Back-Propagation (BP) algorithm is a widely used algorithm for learning weights, that has been extensively employed with respect to many applications domains due to its simplicity. If the accuracy of a generated ANN is not acceptable then the ANN can be trained again using either a different structure or new initial weights [104].

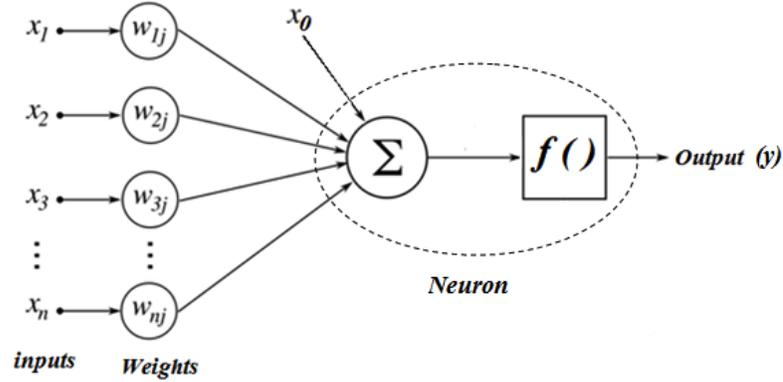


Figure 2.4: Example of a typical perceptron in ANN

Figure 2.4 presents a simple example of an ANN (a perceptron). The perceptron is an early example of an ANN [55, 104]. The  $n$  inputs are given by  $X = x_1, x_2, \dots, x_n$  and each  $x_i$  is connected to the neuron by a weighted link  $w_i$ . Typically, the neuron consists of a summation function along with an activation function (sometimes referred to as the “threshold function”). In a simple case, the output  $y$  will be activated ( $y = 1$ ) if and only if the summation of  $x_0$  and the weighted inputs exceed a threshold value  $t$  as shown in Equations 2.16 and 2.17 and in Figure 2.4.

$$x_0 + \sum_{i=1}^n x_i w_i > t \quad (2.16)$$

$$y = f\left(x_0 + \sum_{i=1}^n x_i w_i\right) \quad (2.17)$$

According to Figure 2.4  $x_0$  is an additional fixed input called the bias neuron which can exist in more than one layers.  $x_0$  is connected to all neurons in the next layer (but not the previous one).  $x_0$  can be set to any value in the activation function for some specific output. The main goal of  $x_0$  is to provide more flexibility and control for the ANN [55, 104]. Beside the simple perceptron, there are many different types of

ANN that have been proposed. A commonly used type of ANN is the Multilayer Feed Forward Neural Network (FNN) where the flow of information is in only one direction (forward). FNN can be seen as an extension of the perceptron with hidden layers and sometimes it is known as a Multi-Layer Perceptron (MLP) [105]. Despite the accurate prediction that can be obtained using supervised ANN, the main limitation of ANN is the large amount of training data required to effectively train them. Moreover, the time complexity for training ANN can be significant which makes them unsuited to mining very large data sets [27].

#### 2.4.5 K-Nearest Neighbours

K-Nearest Neighbour (KNN) is a non-parametric classification technique. KNN was originally proposed by Fix and Hodges in 1951 [43] and is now considered to be one of the most powerful classification techniques available [134]. KNN operates by finding the most similar  $k$  previously labelled records to a new record, and using the knowledge of these pre-existing labels, to label the new record. The main challenges are: (i) the similarity measure to be used, (ii) how to address the situation where the nearest  $k$  records have different labels associated with them and (iii) what the “best” value for  $k$  is. In the case of the work described in this thesis, as will become apparent later in Chapter 6,  $k = 1$  was used thus obviating the need to resolve challenges (ii) and (iii). Records are usually presented using a feature vector representation, in which case similarity between two records can be determined using a simple distance measurements such as the standard Euclidean distance measure (Equation 2.18) or the Manhattan Euclidean distance measure (Equation 2.19). The first is typically used where the distribution is Gaussian, the second where it is Exponential [138].

$$D(x, y) = \sqrt{\sum_i = 1^N (x_i - y_i)^2} \quad (2.18)$$

$$D(x, y) = \sqrt{\sum_i = 1^N |x_i - y_i|} \quad (2.19)$$

The above distance measures require two equal length feature vectors (so that a one to one matching can be achieved). Thus these distance measures are not suited to all types of data and data distributions [138, 103]. Simple distance measures are also not suited to data representations other than feature vector representations. In the context of the work described later in this thesis the KNN algorithm was used in combination with Dynamic Time Warping (DTW) as this allowed for the effective measurement of the similarity between point series [88, 20, 137]. (Note that DTW is discussed more in Chapter 6).

## 2.5 Image Classification in MRI Brain Scan

This section provides a review of some previous related work concerning 3-D MRI brain scan classification. The related work is the work of Shattuck et al. [110], Zhang et al. [140] and Rajini et al. [98].

In [110] a method was proposed to classify “brain tissue” from 3-D MRI brain scans using a partial volume model. A histogram based representation was used to measure brain tissue intensity and noise variance in the image and then classify this into six tissue types using “a posteriori” classifier. Although the result was promising the performance of the classification process was rather slow (2-3 minutes per volume).

Both [140] and [98] proposed similar works based on Discrete Wavelet Transform (DWT) representations. Both proposed a classification method to classify normal and abnormal brains (considering the whole brain) using DWT to extract features from 3-D MR images and then Principle Component Analysis (PCA) to help reduce the number of dimensions. Their results were excellent in terms of both effectiveness and efficiency but it was argued that the difference between normal and abnormal brains might be too obvious.

For work based on Tree based representations, Cocosco et al. [22, 23] proposed an approach to classify “brain tissue” using a minimum spanning tree graph-theoretic approach. This was evaluated using four classes: (i) elderly, (ii) young normal individual brain, (iii) ischemia patients’ brain and (iv) Alzheimer’ brain. Some good results, in terms of effectiveness, were reported. In [82] an approach to classifying MRI brain scans according to different levels of education was considered. The “lateral and third ventricles” of the brain were used and represented using an Oct-tree structure and FSM (SUBDUE algorithm) to generate feature vectors. The classification results were promising in terms of classification effectiveness but the disadvantage was that the run time complexity was high.

## 2.6 Evaluation Criteria

The main aim of the work presented in this thesis was to identify the most appropriate 3-D representation and classification technique in the context of VOIBIC. To identify this representation the proposed representations were evaluated individually and comparatively. More specifically the conducted evaluation was as follows:

1. **Individually:** For each approach (discussed separately in each relevant chapter) the classification accuracy, sensitivity and specificity were used as the effectiveness measures.
2. **Comparatively:** The Friedman and Nemenyi statistical tests were used to demonstrate whether there was a statistically significant difference among the

operation of the proposed approaches.

This section presents an overview of the evaluation measures used with respect to the individual evaluations (classification accuracy, sensitivity and specificity).

The most fundamental mechanism for analysing classifier performance within the data mining community is the confusion matrix as shown in Figure 2.5. For simplicity, the assumption is that each instance can only be assigned one of two classes: “Positive” or “Negative”. Each instance has a known label and a predicted label. Using a confusion matrix shown in Figure 2.5 accuracy, sensitivity and specificity can be calculated using Equations 2.20, 2.21 and 2.22. *Accuracy* is thus the percentage of predictions that are correct. *Sensitivity* is also referred to as “recall” or “true positive rate”. It is measure how well a classifier can be used to classify instances as belonging to a certain class. *Specificity* is also referred to as “recall” or “true positive rate” It is measure how well a classifier can be used to classify instances as belonging to a certain class. *Specificity* corresponds to the “true negative rate” which is similar to Sensitivity but for the negative class label. A good classifier is one that can maximise accuracy, sensitivity and specificity.

		Predicted Label	
		Positive	Negative
Known Label	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Figure 2.5: Example of confusion matrix

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (2.20)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2.21)$$

$$Specificity = \frac{TN}{FP + TN} \quad (2.22)$$

For the comparatively evaluation, details concerning the adopted statistical significance testing are presented later in Chapter 8.

## 2.7 Conclusion

This chapter has presented the background and related work with respect to the work described in this thesis. Previous work regarding image segmentation, image representation and classification have been described with the focus on challenges of MRI brain scan classification. Before any processing can be undertaken it is first necessary to identify the VOI. This requires recourse to image segmentation. Several image segmentation techniques that may be applied to MRI brain scan were described. A review of the (limited) literature concerned with VOIBIC was also presented. In the next chapter the medical MRI brain scan datasets that were used for the purpose of evaluation with respect to the work described in this thesis are described in detail.

## Chapter 3

# Evaluation Datasets

As noted in the introduction to this thesis, the research presented is directed at volumetric data classification. The volumes in this case are Volumes Of Interest (VOI) contained in 3D image data. To act as a focus for the work the classification of 3-D Magnetic Resonance Imaging (MRI) scans of the human brain was considered. The VOI in this case were the left and right (lateral) ventricles. These are cerebrospinal fluid filled spaces at the centre of the brain [79]. Their function is: (i) to act as shock absorbers, (ii) to distribute nutrients to the brain and (iii) remove waste. There are in fact four ventricles in a human brain: two lateral ventricles (referred to as the left and right ventricles), a third smaller ventricle connected to both lateral ventricles and a fourth small ventricle that connects the third ventricle with the spinal cord. Only the left and right lateral ventricles were considered with respect to the focus of the research presented in this thesis. This was because: (i) the lateral ventricles are relatively easy to identify within 3-D MRI brain scans, so facilitating automatic extraction; and (ii) they are much larger than the other two ventricles and consequently can be argued to be more significant.

Two MRI brain scan datasets were used for evaluation purposes with respect to the research presented in this thesis: (i) an Epilepsy dataset (Epilepsy brains v. Healthy brains) and (ii) a Musician dataset (Musician brains v. Non-musician brains). For the Epilepsy dataset, the MRI brain scan volumes were obtained by the Magnetic Resonance and Image Analysis Research Centre (MRIARC), at the University of Liverpool, between the years 1999 and 2004. For the Musician dataset, some of the MRI brain scan volumes were obtained by MRIARC between the years 1999 and 2004, and the remainder by the University of Heidelberg between the years 2000 and 2004. This short chapter provides an overview of these evaluation datasets. Section 3.1 gives details concerning the Epilepsy dataset while Section 3.2 provides details concerning the Musician dataset. The chapter is concluded in Section 3.3 with a summary and a “look ahead”.

### 3.1 Epilepsy Dataset

The epilepsy dataset comprised 210 MRI brain scans. Of these 105 were from healthy people and the remaining 105 from epilepsy patients. Figure 3.1 shows the epilepsy dataset MRI scans grouped according to gender, while Figure 3.2 shows the epilepsy dataset scans grouped by age range. From Figure 3.1, it can be seen that the number of female epilepsy subjects is slightly higher than the number of male subjects, while the number of healthy females is significantly higher than the number of males. From Figure 3.2 it can be seen that the distribution of age for epilepsy subjects is normal although the overall range of ages is less than that for healthy subjects. The distribution for healthy subjects is more balanced, while that for epilepsy subjects is not. The percentages of males and females in each age group, and the percentages of Epilepsy and non-Epilepsy subjects in each age group, is shown in Figure 3.3. An example of an MRI brain scan from an epilepsy patient is given in Figure 3.4

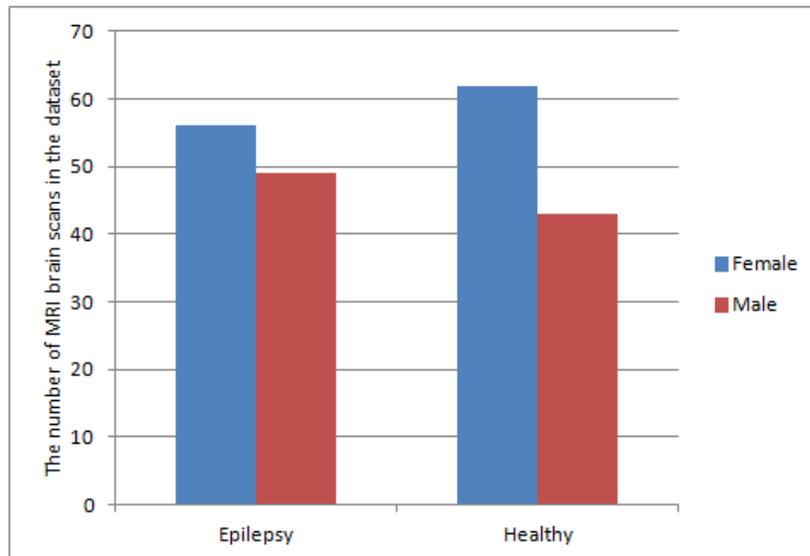


Figure 3.1: Epilepsy MRI brain scan data grouped by gender

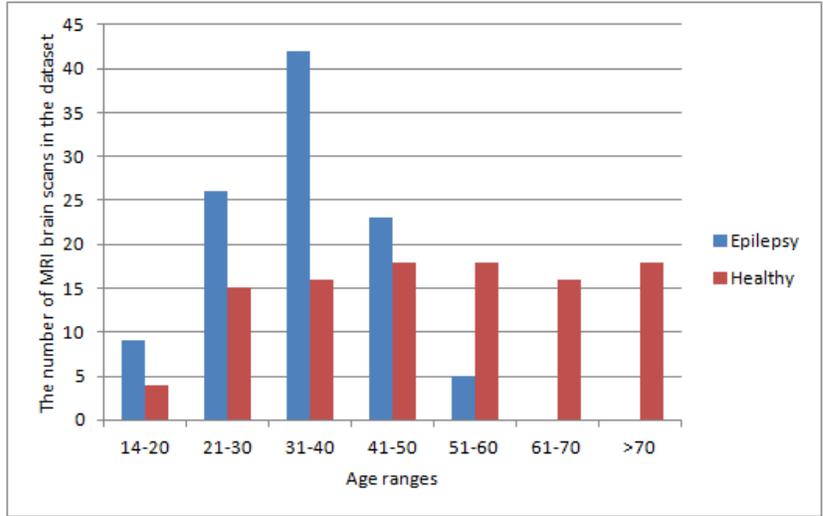


Figure 3.2: Epilepsy MRI brain scan data grouped by age range

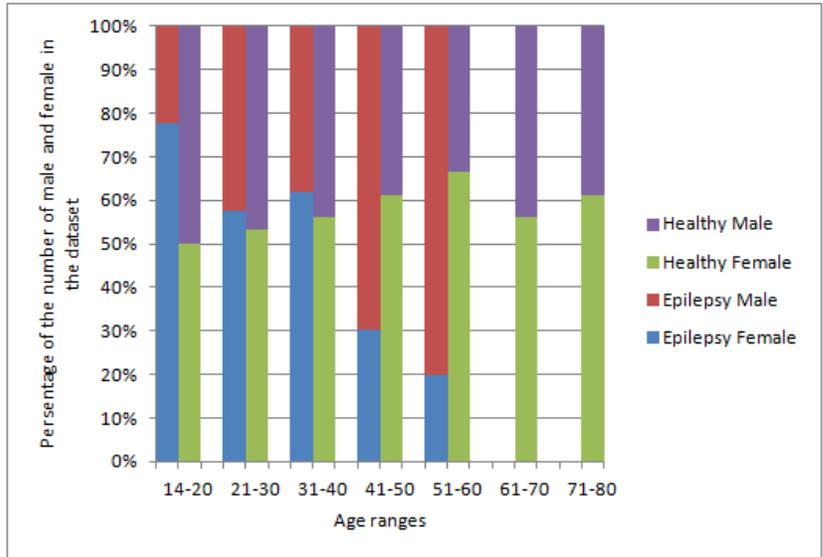


Figure 3.3: The percentages of male and female subjects grouped by age, and the percentages of the Epilepsy and non-Epilepsy subjects grouped by age, for the Epilepsy MRI brain scan data

Although the epilepsy data is not equally distributed in terms of gender and age it was considered that this would not adversely effect the evaluation results obtained because the evaluation was directed at the nature of the proposed algorithms and techniques for predicting class labels to be associated with 3-D objects. The intention was not to investigate the nature of epilepsy in the context of age and/or gender. However, the work described in this thesis is also concerned with determining whether data augmentation will have a beneficial effect with respect to classification accuracy (research question 7 presented in Section 1.2 in Chapter 1). Therefore, as will become clear later in this thesis, individual representations were evaluated with and without

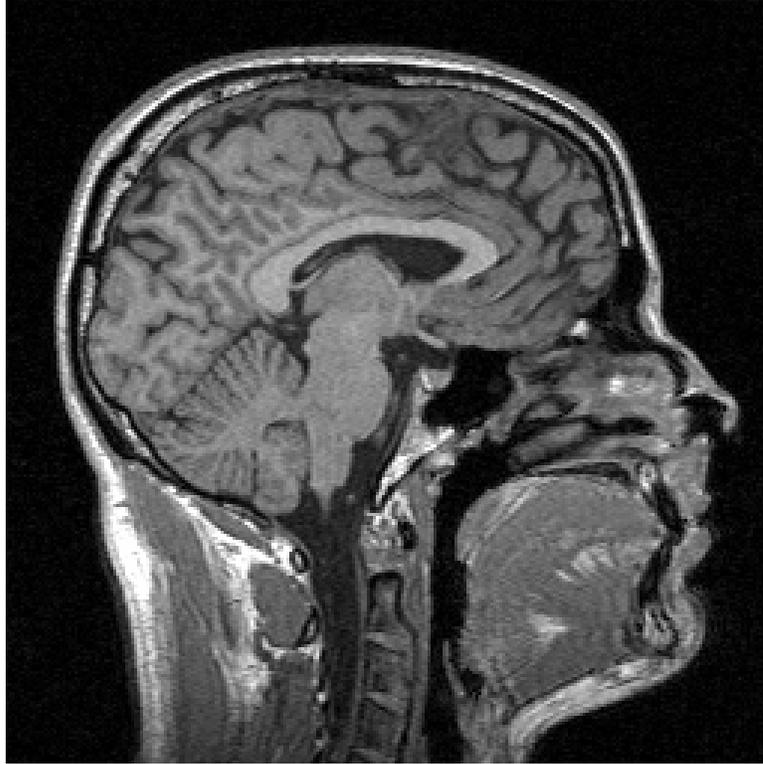


Figure 3.4: Example of MRI brain scan from an Epilepsy patient

augmentation. In each case the augmentation comprised the inclusion of age and gender attributes.

It should also be noted here that there have been a number of previous studies where data mining has been applied to epilepsy MRI scan data. Of note with respect to this thesis is [40] and [37]. The significance of these two publications is that, although only 2-D representations were considered, the same epilepsy MRI brain scan data set was used as that used in this thesis. Therefore, a comparison of the results obtained with respect to the 3-D approaches proposed in this thesis and those obtained with respect to the 2-D approaches presented in [40] and [37] can be made.

## 3.2 Musician Dataset

The musician dataset comprised a total of 160 MRI brain scans. Of these 80 were from musicians and the remaining 80 from non-musicians. The musician brain scans were collected from members of: (i) the Royal Liverpool Philharmonic Orchestra, (ii) the Heidelberg Music Conservatory, (iii) the Frankfurt Am Main Orchestra and (iv) the Berlin Philharmonic Orchestra. The MRI brain scan volumes for the Royal Liverpool Philharmonic Orchestra were scanned at MRIARC between 1999 and 20004, and the rest were scanned at the University of Heidelberg between 2000 and 2004. Figure 3.5 shows the Musician MRI scan data grouped according to gender, while Figure 3.6

shows the Musician MRI brain scan data grouped according to age range. From Figure 3.5 it can be seen that the number of female musicians was significantly lower than the number of male musicians, while in the case of female non-musicians the gender distribution was more balanced. From Figure 3.6 it can be seen that the majority of musicians considered were in the age ranges of ‘21-30’ and ‘31-40’. There were no musicians older than ‘70’. On the other hand, the distribution according to age range for non-musicians was relatively balanced. The percentages of males and females in each age range, and the percentages of musicians and non-musicians in each age group, is given in Figure 3.7. An example of MRI brain scan from a musician is given in Figure 3.8

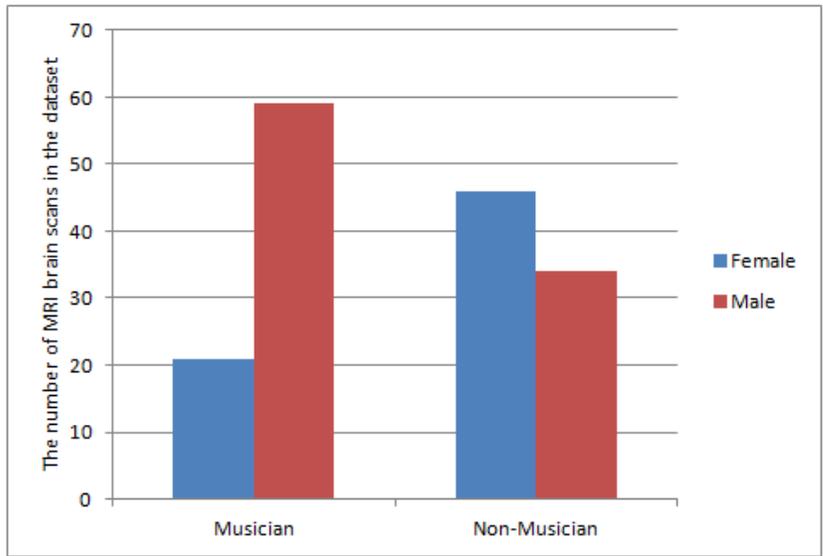


Figure 3.5: Musician MRI brain scan data grouped by gender

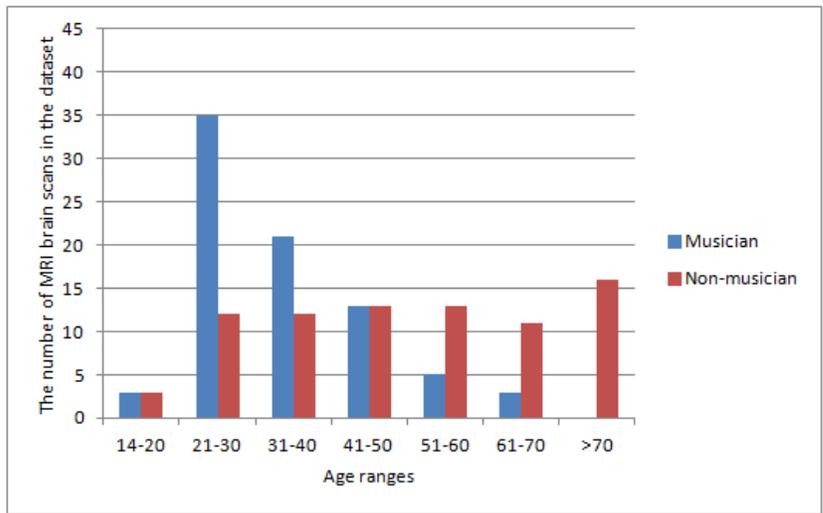


Figure 3.6: Musician MRI brain scan data grouped by age range

The above referenced figures indicate that, as in the case of the epilepsy data set, the musicians dataset was also unbalanced in terms of gender and age. However, again as in the case of the epilepsy data set, this was deemed not to be significant as the objective of the evaluations using the musicians data was the analyse of the performance of the proposed algorithms and techniques in terms of classification accuracy. The objective was not to analyse the musicians data set so as to attempt to draw some conclusions about musical ability in the context of age and/or gender. However, as in the case of the Epilepsy data set, experiments reported on later in this thesis, were conducted using the proposed representations augmented with age and gender information.

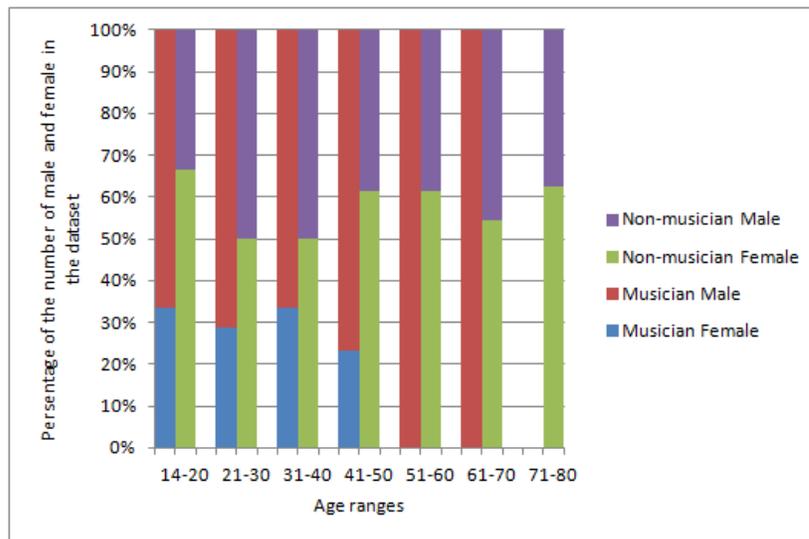


Figure 3.7: The percentages of male and female subjects grouped by age, and the percentages of musicians and non-musicians grouped by age, for the Musician MRI brain scan data

There have been a number of previous studies where data mining has been applied to the musicians MRI scan data. Of note with respect to this thesis is again the work of [40] and [37] (who also did work on the epilepsy data set). The significance is that the musician MRI scan data used in [40] and [37] is again the same as that used with respect to the research presented in this thesis. Therefore comparisons can be made with respect to the results obtained using the approaches proposed in this thesis and the results obtained using the approaches presented in [40] and [37] (although the latter considered the musicians data only in terms of 2-D).

### 3.3 Conclusion

This short chapter has introduced the two 3-D MRI brain scan datasets used with respect to the evaluation of the work reported later in this thesis. The datasets were precisely balanced with respect to the class labels of interest, musician versus non-



Figure 3.8: Example of MRI brain scan from a musician

musician and epilepsy versus non-epilepsy. The datasets were less balanced with respect to gender and age range distribution, however, it was conjectured that this would not adversely effect the evaluation results produced with respect to the quality of the techniques considered (although the evaluations reported on later in this thesis include results obtained by augmenting the proposed representations with age and gender meta data). The next chapter considers the segmentation of the lateral ventricles from 3-D MRI brain scan data, a necessary precursor to any consideration of representation mechanisms to support the eventual desired classification.

## Chapter 4

# Volume of Interest Identification

The broad research domain for the work presented in this thesis, as already noted, is volumetric data classification. The Volumes Of Interest (VOI) are 3-D objects contained in 3-D images (volumes). For the volumetric classification to operate the volumes need to be represented in such a way that classification techniques can be applied. One of the objectives of the research presented in this thesis was the study of different ways to represent certain 3-D objects, in way that facilitate a subsequent classification process. The key point of VOI representation is to exclude those elements which will not contribute to the effectiveness of any subsequent analysis, while retaining those that will. However, before this can be done the VOI need to be extracted from the input volumes, a process known as *segmentation*.

In the context of 2-D data the identification of Regions Of Interest (ROIs) has been well researched; it is an important step in 2-D image analysis of all kinds [16], including 2-D image classification. The reason is that the accuracy with which the nature of the ROI is captured directly affects the effectiveness of any subsequent analysis (such as classification or prediction). The identification of VOI has been less well studied. One reason is that ROI identification is computationally cheaper, in terms of both storage cost and processing time, than ROI identification, especially when the volumes under consideration are large.

As established in previous chapters, the application focus for the research presented in this thesis 3-D MRI scans of the human brain. The VOI in this case are the left and right lateral ventricles that can be observed in 3-D MRI brain scans. The identification of the left and right ventricles in MRI brain scan data is thus the start point for the work presented in this thesis. As noted above, a VOI is simply the 3-D extension of the more commonly found concept of ROI. Techniques used for ROI identification (see Chapter 2) can clearly be extended to VOI identification.

An example of a 3-D MRI brain scan is given as Figure 4.1 where the “lateral ventricles” are the dark areas at the centre of the brain. Note that 3-D MRIs comprise a sequence of two dimensional (2-D) “slices” through the brain in each of the three

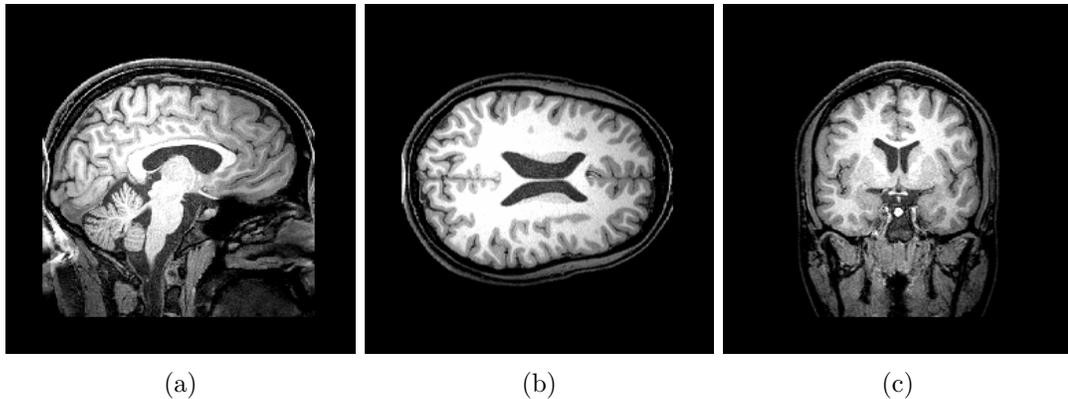


Figure 4.1: Example of a 3-D brain MRI scan: (a) Sagittal (SAG) plane, (b) Transverse (TRA) plane and (c) Coronal (COR) plane

cardinal planes: (i) Sagittal - SAG (left to right), (ii) Coronal - COR (front to back) and (iii) Transverse - TRA (top to bottom). Recall from the previous chapter that the MRI brain scan volumes used to illustrate the contents of this chapter, and with respect to the evaluation of the proposed 3-D segmentation techniques presented in this chapter, were obtained from: (i) the Magnetic Resonance and Image Analysis Research Centre (MRIARC) at the University of Liverpool and (ii) the University of Heidelberg; during the years 1999-2004. Recall also that each brain scan is composed of 256 two dimensional parallel image slices in each plane. The resolution of each image slice is  $256 \times 256$  pixels with colour defined using an 8-bit grey scale (thus 256 colours).

The work presented in this chapter is concerned with the identification of VOI, specifically the left and right lateral ventricles found in 3-D MRI brain scan data. As noted in Chapter 2 there were a number of “off-the-shelf” 3-D segmentation options that could have been adopted; however, as also noted in Section 2.2 in Chapter 2, these were found to be unsuitable with respect to the 3-D MRI brain scan data of interest with respect to this thesis in that they did not produce a satisfactory result. Instead two alternative bespoke algorithms for extracting the lateral ventricles (the VOIs of interest) from MRI volumes are proposed in this chapter: (i) Volume Growing and (ii) Bounding Box (early versions of these algorithms were published in [122]). At a high level both algorithms work in a similar manner. In each case the input to the process is a set of slices, in a particular dimension, for a given 3-D MRI brain scan. Prior to the application of the algorithms the given MRI brain scan volume is first preprocessed. The preprocessing comprises: (i) slice capture and registration and (ii) contrast enhancement. One or other of the proposed algorithms are then applied; the output is a set of voxels representing the VOI (the lateral ventricles). To evaluate the effectiveness of the proposed segmentation algorithms a Bland-Altman analysis [6] was applied.

All the procedures that were either adopted or developed are fully described in this

chapter which is organised as follows. In Section 4.1 the required image preprocessing is described. Next, the Volume Growing and Bounding Box algorithms are explained in Section 4.2. The evaluation of the results obtained using the proposed algorithms is presented and discussed in Section 4.3. Finally, the chapter is concluded in Section 4.4 with a summary of the main findings and a “look ahead”.

## 4.1 Image Preprocessing

As noted in Chapter 2, often pre-processing is an essential precursor to image analysis of any kind, regardless of whether the images are 2-D or 3-D. With respect to the work presented in this thesis, the preprocessing of the MRI brain scan volumes comprised: (i) slice capture and registration and (ii) contrast enhancement. These are considered in further detail in the following two subsections.

### 4.1.1 Slice Capture and Registration

Each MRI brain scan volume, making up the 3-D MRI datasets used with respect to the work presented in this thesis, were in two files: (i) an image file (extension “.img”) holding the actual data in a binary format and (ii) a header file (extension “.hdr”) which contains information about the data such as voxel size and number of voxels in each dimension [33]. There are a number of software tools which are compatible with the “.hdr” and “.img” file formats that can be used to view 3-D MRI volumes. For the work described in this thesis the MRICro [1] 3-D file viewer and analysis software system was used to view images in the “.hdr” and “.img” file formats and transform them into Portable Network Graphics (“.png”) image files. After capturing all image slices, using the MRICro software, a registration process was applied. The purpose of the registration process was to ensure that all image slices conformed to the same reference framework. With respect to the work presented in this thesis the alignment (registration) of image slices was conducted semi-manually, within MRICro, in order to ensure that the image slices conformed to a single axes of orientation. For each 3-D MRI brain scan, there are 768 image slices, 256 for each (SAG, COR and TRA) plane. The image slices in each plane might not be in the same axes. Thus they needed to be adjusted manually, once for each plane. Note that all image slices in the same plane are already in the same alignment. Then after an image slice in a plane has been adjusted manually, the rest of the image slices in that plane can be adjusted automatically using the same angle of adjustment. The processing time required was approximately 1 to 2 minutes per brain scan. An example of the MRICro software interface, showing a number of brain image slices registered in the context of the three cardinal dimensions, is shown in Figure 4.2.

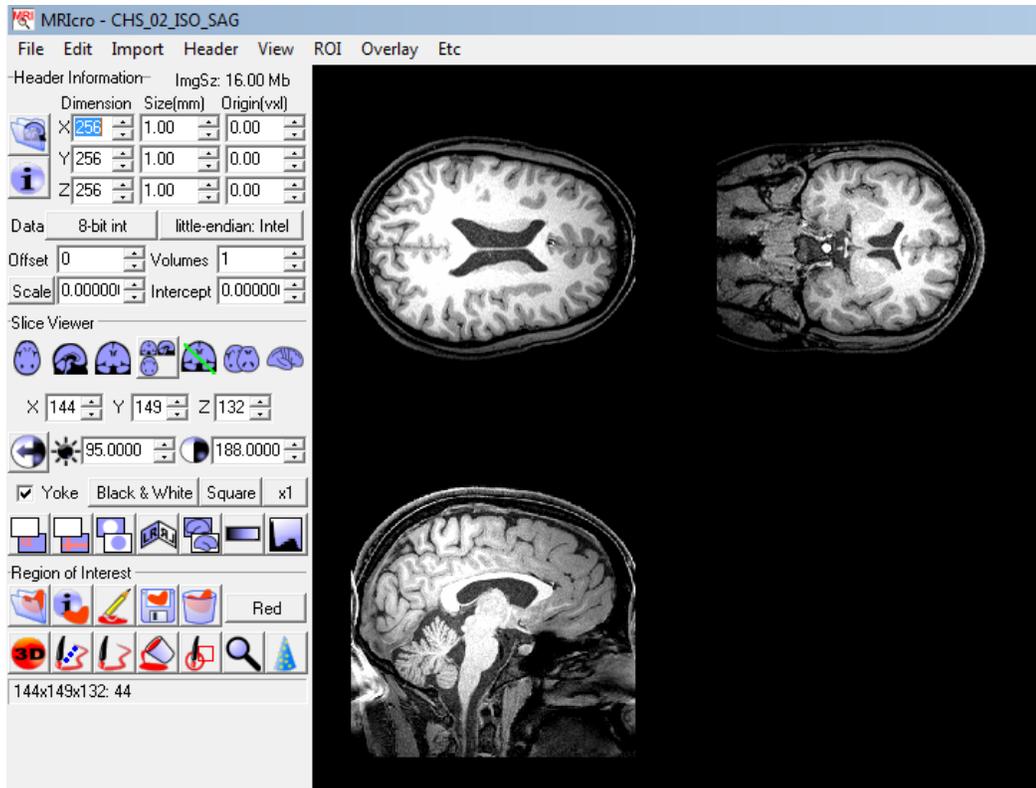


Figure 4.2: Example of the MRICro software interface showing three different views of a MRI brain scan image

#### 4.1.2 Contrast Enhancement

The objective of contrast enhancement is to clarify the boundaries of the VOI (lateral ventricles) by considering the brightness of pixels. A thresholding technique [89] was used for this purpose. This was selected because this technique is considered to be the most effective when an objects colours are obviously different from its background colours [33]. Note that thresholding is also applicable where object colour and background colour are not noticeably distinguishable by humans, but can be distinguished when a brightness range is used. Figure 4.3 shows a number of MRI brain scan slices. In the figure the lateral ventricles are represented by the dark areas towards the middle of the brain, surrounded by brain tissue which appears as grey (or white) matter. Generally, the contrast between the ventricle and other parts of the brain is easily noticeable, but in some slices (such as SAG slice number 160 in Figure 4.3), it is difficult to identify the boundary of the ventricle because there is a grey shade (which is brain tissue) appearing in the dark area. In this case, the thresholding technique will enhance the contrast so as to aid the identification of the ventricle. During thresholding, each pixels brightness is compared with a predefined threshold. It will be considered to be part of the object of interest or not according to this threshold. If the pixel value is greater than the threshold the pixel colour is set to some predefined distinguishing

colour (black in the case of the work described here), otherwise it will be identified as background and set to an alternative predefined colour (white in the case of the work described here).

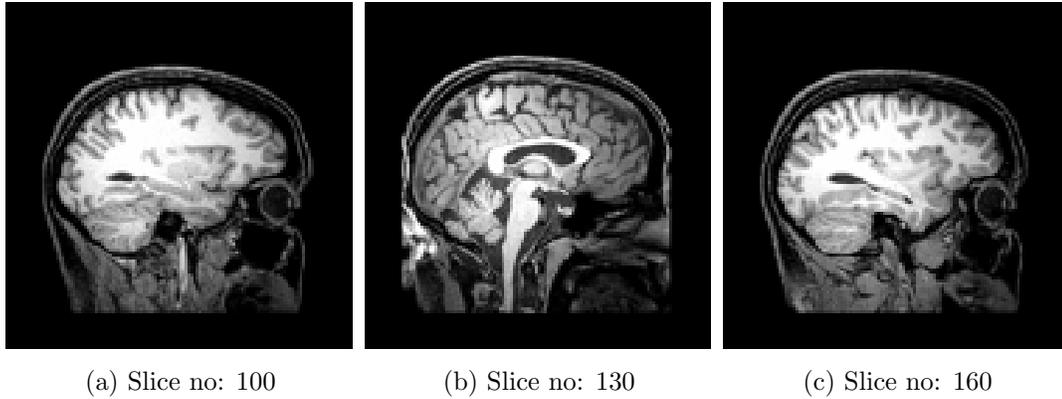


Figure 4.3: Examples of 3-D MRI brain scan slices in the Sagittal (SAG) plane

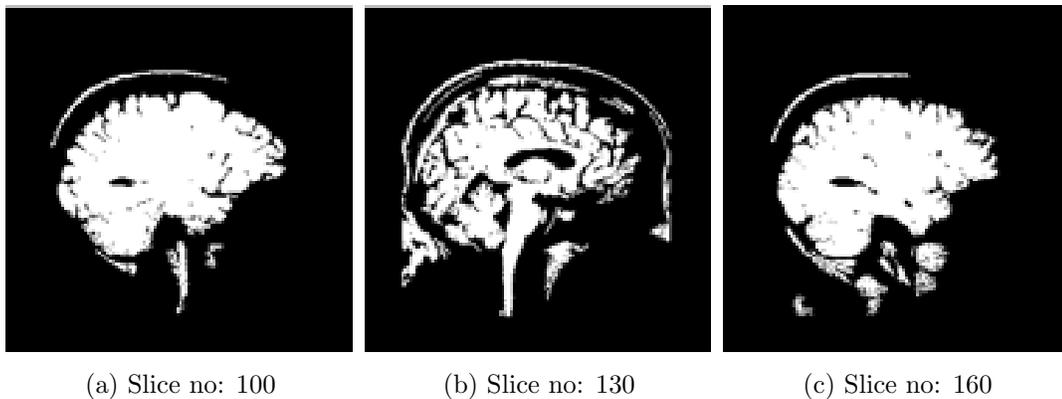


Figure 4.4: Examples of 3-D MRI brain scan slices in the Sagittal (SAG) plane after applying thresholding (threshold value = 0.30)

The key to the success of the thresholding process is thus the selection of a most appropriate threshold value. For identifying ventricles in the 3-D MRI volumes of interest the threshold value was manually set to 0.30. This was selected after the effect of a range of different threshold values had been observed by a domain expert<sup>1</sup>. As a result, if a pixel was darker than 0.30 (of interest) it was set to black, and otherwise (not of interest) to white. Figure 4.4 shows the same MRI brain scan volume slices as those given in Figure 4.3, but after application of the thresholding technique. Note that for the work described in this thesis the thresholding function provided by Matlab [2] was used (Matlab provides a suite of image processing functions). Using Matlab the threshold value can be automatically assigned by the software or manually set by a human user; the later was adopted.

<sup>1</sup>Dr. Vanessa Sluming, a leading Neuroimaging Scientist at University of Liverpool.

## 4.2 Image Segmentation

As already noted, to identify VOI in 3-D data requires the application of some form of segmentation process. As noted in Chapter 2, possible options were discussed in detail in Section 2.2. In the context of 2-D data many segmentation algorithms exist, popular examples include: (i) Otsu’s Thresholding method [93], (ii) Adaptive Region Growing method [84] and (iii) Deformable and Level Set methods [74, 108]. However, it was found that application of these “off-the-shelf” methods did not produce a satisfactory result, possibly because the colour of the ventricles and their surrounding areas (brain tissue, and third and fourth ventricles) are very similar and connected to each other in some views and the time complexity of the segmentation was expensive. Instead, two alternative 3-D segmentation algorithms were proposed directed specifically at segmenting objects (such as the lateral ventricles) in MRI brain scan volumes: (i) volume growing and (ii) bounding box. Details concerning these algorithms are presented in detail in the following two subsections.

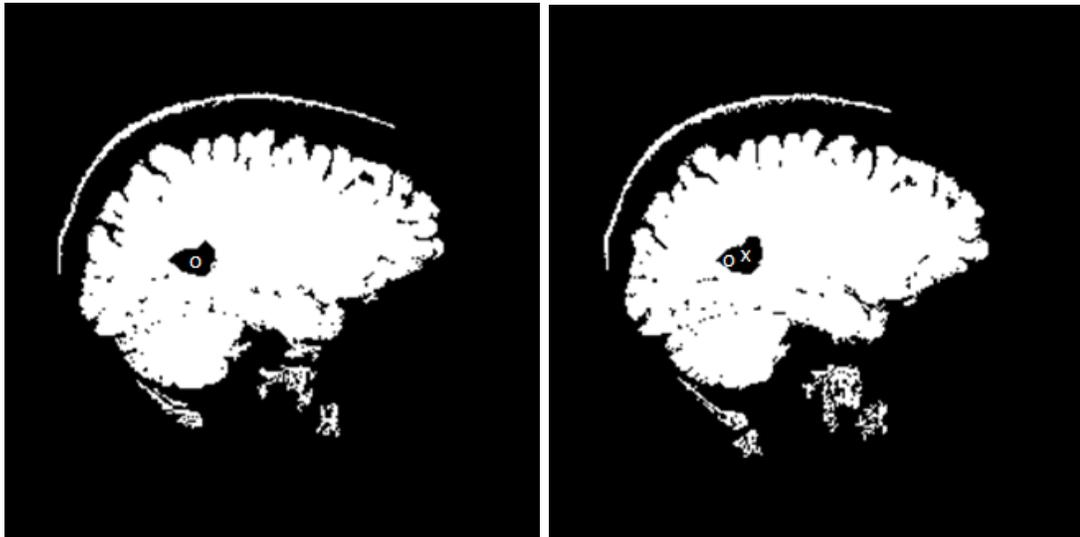
### 4.2.1 3-D Segmentation Using The Volume Growing Algorithm

The Volume Growing 3-D Segmentation algorithm comprises the following steps:

1. Select the starting slice.
2. Manually identify a “guide point” in the VOI (black area).
3. Grow the (2-D) regions with respect to the current slice.
4. Recalculate the guide point with respect to the region identified in (3).
5. Move to the slices on either side of the current slice if at start, otherwise move forward or backward from the start slice as appropriate.
6. If the guide point is in a white region stop, otherwise go to (3).

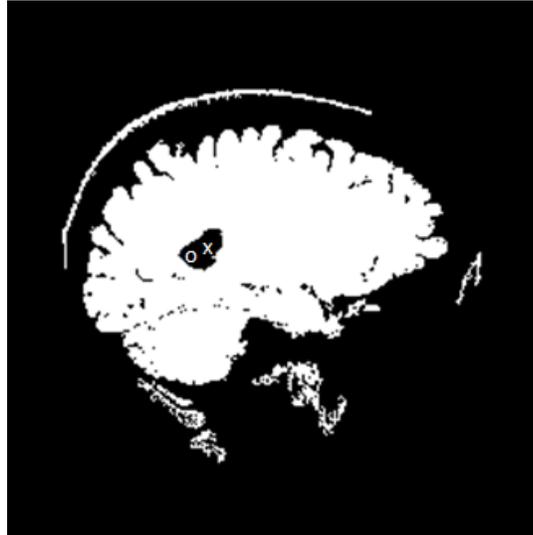
The Volume Growing algorithm process requires the starting slice and the guide point to be manually identified (Steps 1 and 2). The guide point is the pixel (voxel) location in the 2-D MRI slice from which the region growing is to commence. Once the guide point and start slice have been selected the region growing can start (Step 3). This is an automated process. We move out from the guide point in all possible directions in the start slice, from the guide point, by first considering the immediate neighbours to the guide point, then the neighbours of the neighbours, and so on. As we proceed we reject pixels representing white space and store pixels representing black space. We only grow black pixels. The process continues until the region cannot be grown any further (we have only white pixels). Once the region of interest within the current slice has been identified the location of the guide point is recalculated (Step

4) by identifying the geometric centre of the region. This will then be the guide point used when the process moves to the neighbouring slice(s). We either move to the two neighbouring slices on either side of the current slice only if at the start of the process, or simply the next slice once the process has started and repeat the process (Step 5). The algorithm does this until guide points for the next slices are arrived at that do not represent a black pixel. In this manner the volume is grown and all the ventricular pixels are collected. Note that the process is repeated twice; once for the left ventricle and once for the right ventricle.



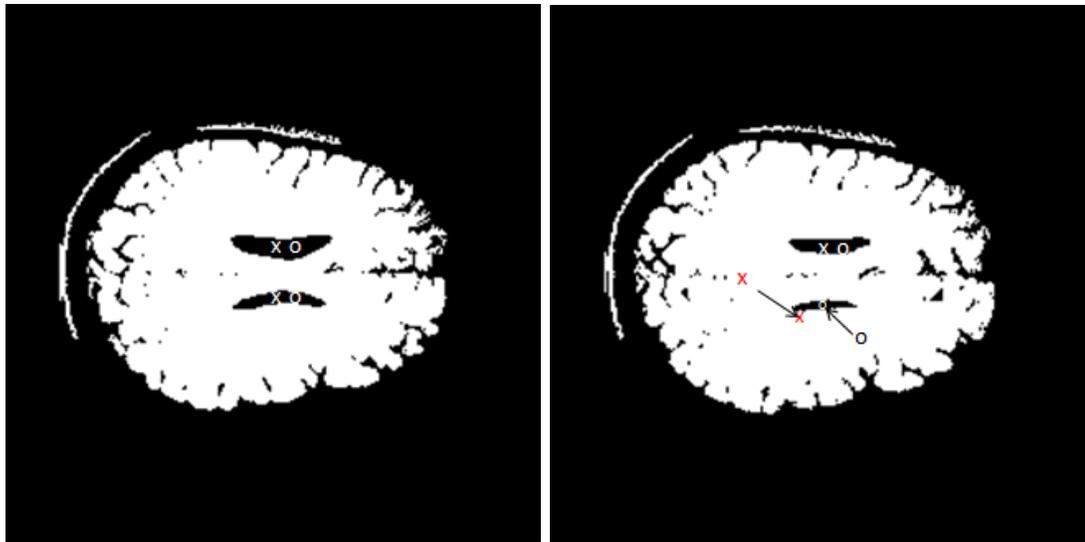
(a) SAG slice no: 106

(b) SAG slice no: 107



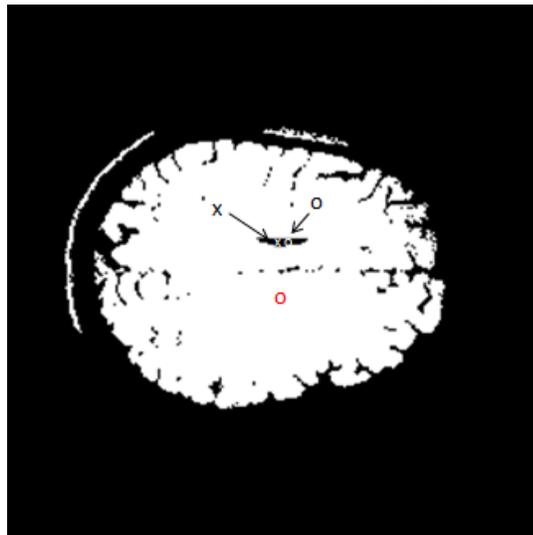
(c) SAG slice no: 108

Figure 4.5: Illustration of the Volume Growing 3-D Segmentation Algorithm in the Sagittal (SAG) plane



(a) TRA slice no: 159

(b) TRA slice no: 160



(c) TRA slice no: 161

Figure 4.6: Illustration of the Volume Growing 3-D Segmentation Algorithm in the Transverse (TRA) plane

Figures 4.5 and 4.6 illustrate the Volume Growing segmentation algorithm. With respect to the figures image slices 106, 107, and 108 are in the SAG plane; and slices 159, 160, and 161 in the TRA plane. In each case the guide point for the volume growing is shown with an 'o', while the recalculated guide point is shown with an 'x'. Thus in the SAG plane (Figure 4.5), and assuming SAG slice 106 is the start slice, the growing proceeds from this manually identified guide point. Once the ventricle voxels in this slice have been collected (grown) the guide point is recalculate ('x'). We then move to the slices on either side, 105 (not shown in Figure 4.5) and 107. In slice 107 the recalculated guide point is indicated with an ('x'), and so on. The algorithm works iteratively and is terminated when the condition in step 6 is reached. In the case of

TRA slice 160 in Figure 4.6, where the guide point is not in a “black area” (the red ‘x’), the process collects the ventricle voxels and then terminates with respect to this branch of the volume growing. Therefore, slice 161 will not be processed. Note that both the left and right ventricles are processed in this manner.

A disadvantage of the Volume Growing algorithm is the need to manually select a start slice and guide point. This is because the colour of the ventricles and their surrounding areas are similar, or exactly the same in some cases except for a thin brighter boundary line. Thus it is difficult to automatically calculate the guide point, especially in the case of “small brains” where it is possible that an automatically calculated guide point may be wrongly calculated and the algorithm will therefore perform unsuccessfully.

### 4.2.2 3-D Segmentation Using The Bounding Box Algorithm

The proposed Bounding Box 3-D Segmentation algorithm comprises three steps:

1. Define a bounding “box” that is expected to encompass the ventricles of interest with respect to all relevant slices in the MRI volume.
2. For each slice collect the black pixels (voxels).
3. Apply noise removal.

The Bounding Box segmentation algorithm initially requires the specification of a pre-defined target area (bounding box). The bounding box is rectangular in shape and defined by the coordinates of its eight corners. To ensure that the bounding box encompasses all the VOI voxels it needs to be defined in such a manner that it is considerably larger than the expected VOI area (so that nothing is missed). Some examples are given, in the context of ventricle segmentation, in Figure 4.7. In the following step (Step 2) all black pixels are collected from each slice that is located within the bounding box. Because the bounding box is defined so that a considerably larger area than the expected ventricle area is covered, some black pixels located outside the ventricle area (noise pixels) will also be collected. These are removed in Step 3 using a simple noise reduction technique whereby the black pixels that are not connected to the largest group of connected pixels were removed. In other words the largest group of pixels is assumed to represent the VOI. Note that the reason that the noise reduction was not applied in the case of the Volume Growing algorithm was that this algorithm operated from “inside to out”, thus all identified voxels were deemed to be relevant (connected to the VOI).

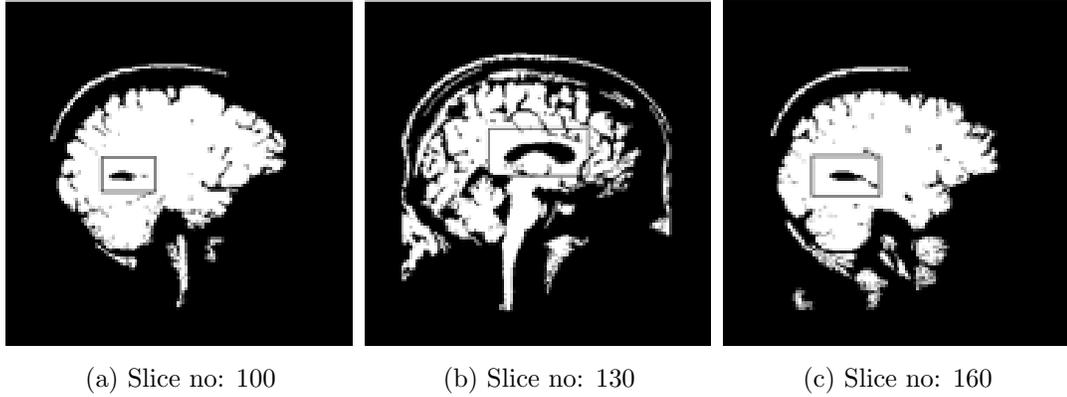


Figure 4.7: Pre-defined target areas (Bounding Boxes) required for the Bounding Box 3-D segmentation algorithm with respect to MRI brain scan slices in the Sagittal (SAG) plane

### 4.3 Evaluation

According to the previous subsections (Sub-sections 4.2.1 and 4.2.2) the output of the proposed segmentation algorithms is a set of coordinate points ( $x$ ,  $y$  and  $z$ ) in the three dimensional Cartesian coordinate system (thus point clouds). The point clouds can be transformed into a 3-D image by using various existing software products. An example of an extracted lateral right ventricle 3-D image generated using the Meshlab open source software [3] is shown in Figure 4.8.

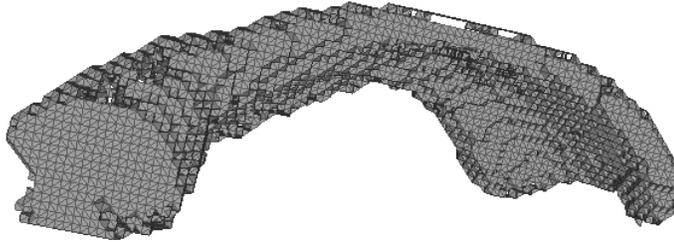


Figure 4.8: Example of an extracted right ventricle generated using Meshlab

To evaluate the proposed 3-D segmentation algorithms experiments were conducted using a volumetric data set comprising 85 MRI 3-D brain scans (of healthy subjects) held by the Magnetic Resonance And Image Analysis Research Centre (MARIARC) at the University of Liverpool. All of the images were a subset of the Epilepsy dataset introduced in Chapter 3. The reason that this subset of images was selected for evaluation purpose was because the volume sizes of these 85 MRI brain scan had been manually calculated by a domain expert<sup>2</sup>. The process of manual calculation of the ventricle volumes was conducted by manually marking pixels which are in the ventricle

<sup>2</sup>Dr. Vanessa Sluming, a leading Neuroimaging Scientist at University of Liverpool.

area on a computer screen, slice by slice in the COR plane. The number of marked pixel on each slice was then summed and considered to represent volume of the given ventricle in terms of voxels. As noted in Section 4.1.1, for each NRI brain scan there are 256 image slices (256 x 256 pixels per slice) in each plane, the manual process of determining volumes was therefore very time consuming, a few hours per brain scan. By applying the proposed algorithms, using two sampling planes (SAG and TRA), to both the left and right ventricle, eight volumes were obtained: (i) Volume Growing SAG plane (left ventricle), (ii) Volume Growing SAG plane (right ventricle), (iii) Volume Growing TRA plane (left ventricle), (iv) Volume Growing TRA plane (right ventricle), (v) Bounding Box SAG plane (left ventricle), (vi) Bounding Box SAG plane (right ventricle), (vii) Bounding Box TRA plane (left ventricle), and (viii) Bounding Box TRA plane (right ventricle). The obtained volume sizes ( $\text{mm}^3$ ) were then compared with the manually calculated volume sizes ( $\text{mm}^3$ ) obtained by the domain expert.

The reason why the COR plane was not used was that the lateral ventricles and their surrounding area were sometimes connected to each other in this view (had the same colour). This is illustrated in Figure 4.9 where it can be seen that the area of the lateral ventricles (left and right dark areas in the middle of the figure) connect to the area representing the third ventricle (dark area in the figure below the lateral ventricles). It was therefore not possible to extract the boundary of the ventricles using this plane.

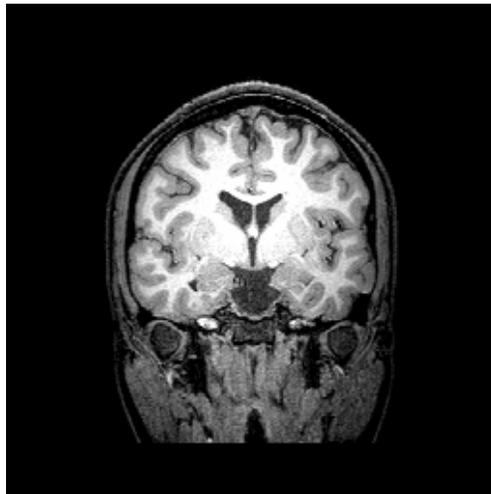
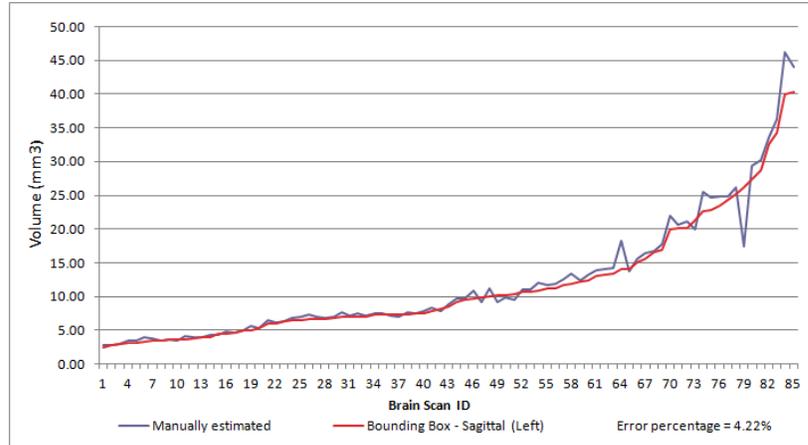


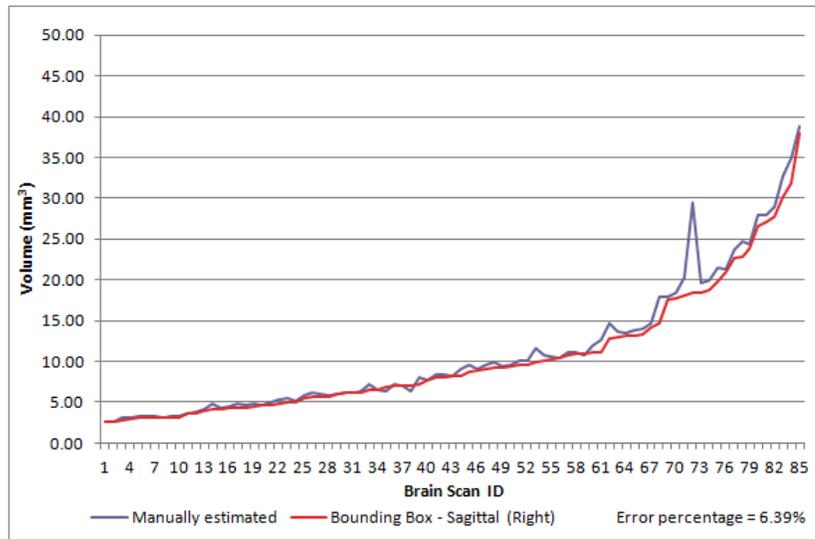
Figure 4.9: Example of a MRI brain scan - COR plane

The results of the comparison are presented in Figures 4.10 to 4.13, where the volumes obtained using both 3-D segmentation algorithms, and the manual technique, are plotted. Figures 4.10 and 4.11 show the results in the SAG and TRA planes, when using the Bounding Box 3-D segmentation algorithm, while Figures 4.12 and 4.13 show the results in the SAG and TRA planes, when using the Volume Growing 3-D segmentation algorithm. In each case the results for both the left and right ventricles

are given as two separate plots. For each plot the identification number for each volume, sorted according to the associated ventricle size, is listed along the x-axis, while the y-axis gives the volume size in  $\text{mm}^3$ .

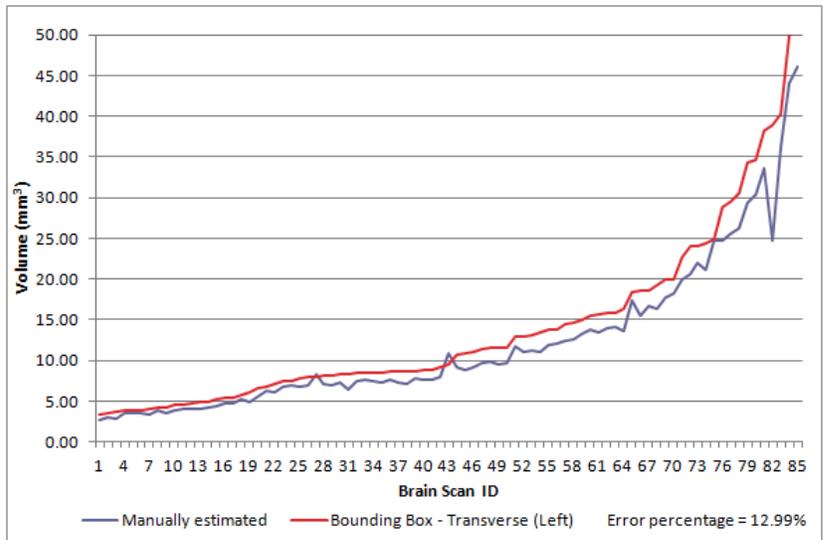


(a) Left Ventricle

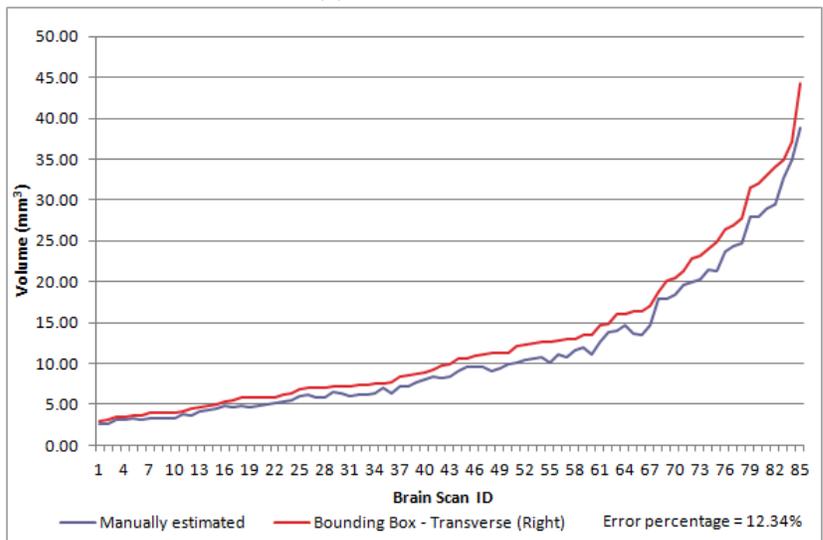


(b) Right Ventricle

Figure 4.10: Comparisons between volumes obtained using manual estimation and the Bounding Box algorithm in the Sagittal (SAG) plane

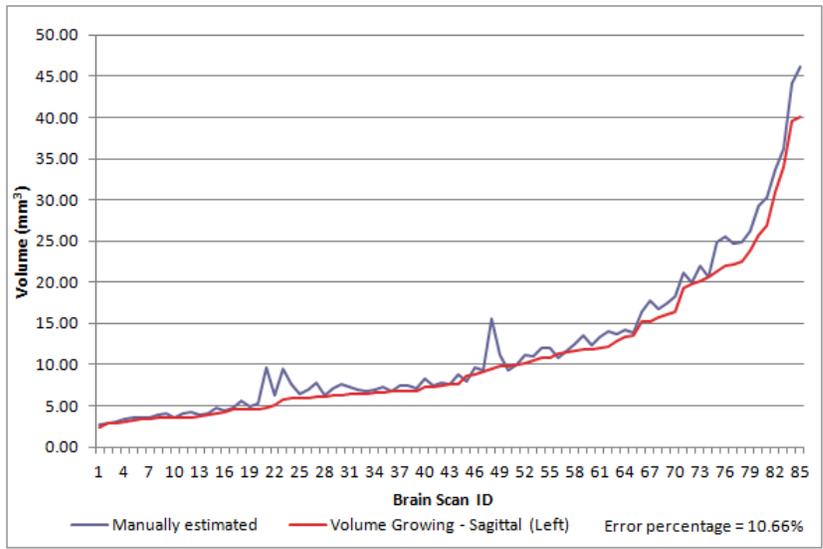


(a) Left Ventricle

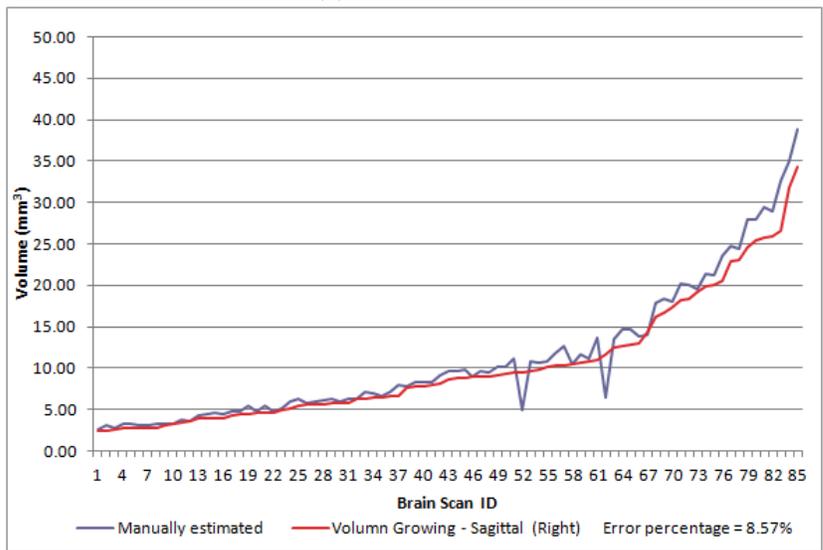


(b) Right Ventricle

Figure 4.11: Comparisons between volumes obtained using manual estimation and the Bounding Box algorithm in the Transverse (TRA) plane

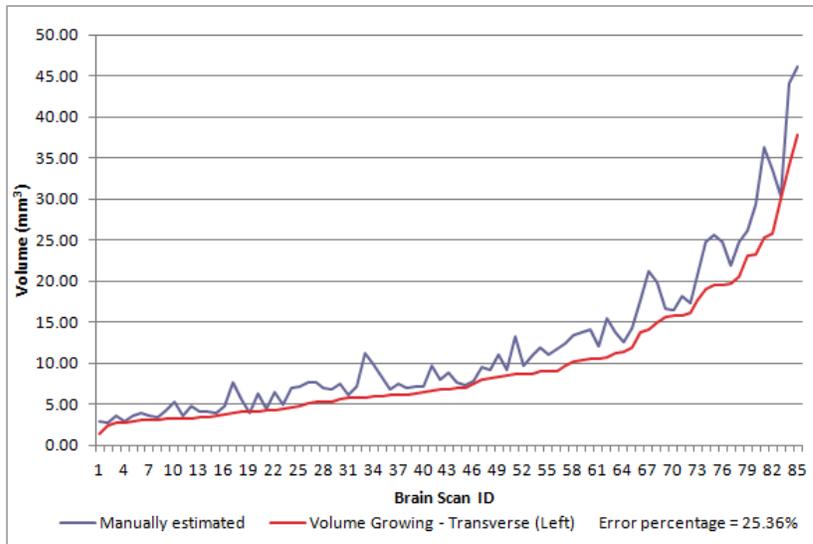


(a) Left Ventricle

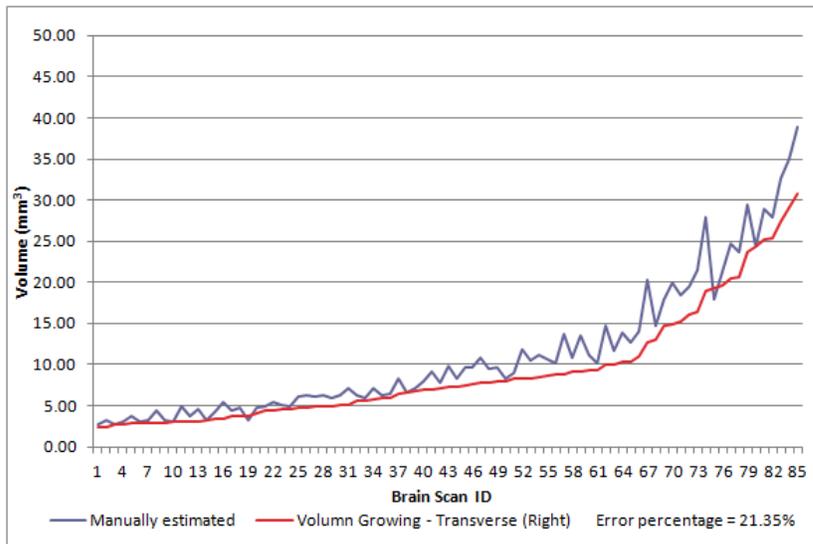


(b) Right Ventricle

Figure 4.12: Comparisons between volumes obtained using manual estimation and the Volume Growing algorithm in the Sagittal (SAG) plane



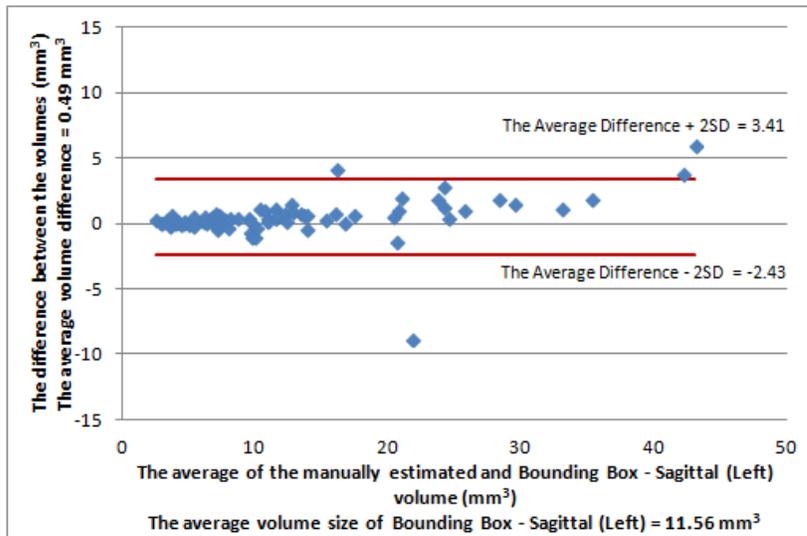
(a) Left Ventricle



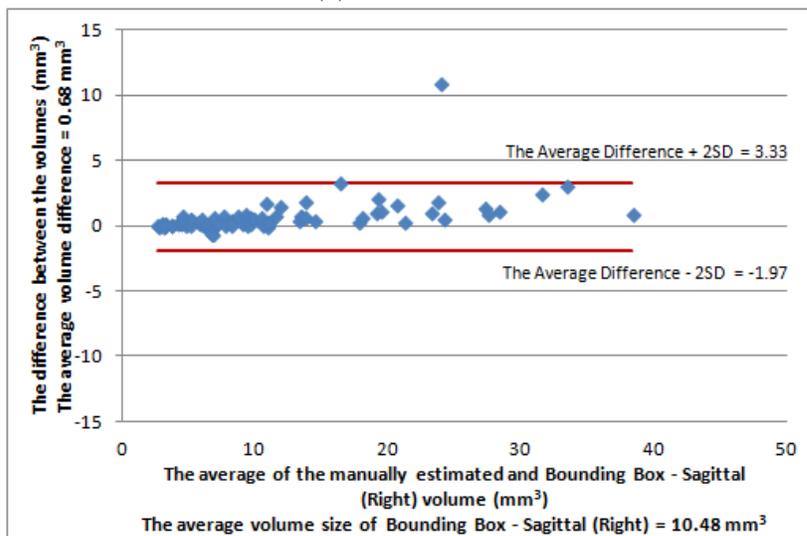
(b) Right Ventricle

Figure 4.13: Comparisons between volumes obtained using manual estimation and the Volume Growing algorithm in the Transverse (TRA) plane

From the Figure 4.10 to 4.13 it can be seen that all the segmented volume values were marginally different to the manually estimated volume values. Of course the manually estimated volumes do not provide a “gold standard”, and may themselves be flawed due to human error. However they did provide a benchmark. Closer inspection of the figures considering the error percentages shown in Figure 4.10 to 4.13 indicates that the best performing algorithm, in comparison with the manual technique, was found to be the Bounding Box algorithm when used in the SAG plane.

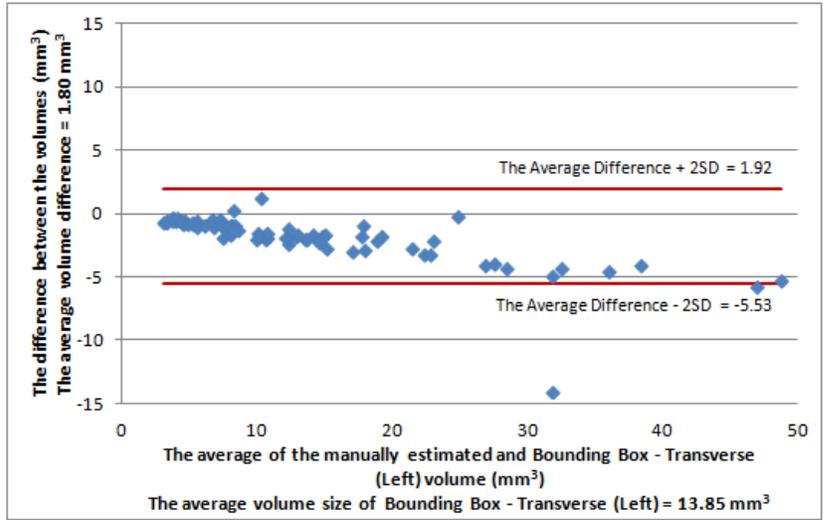


(a) Left Ventricle

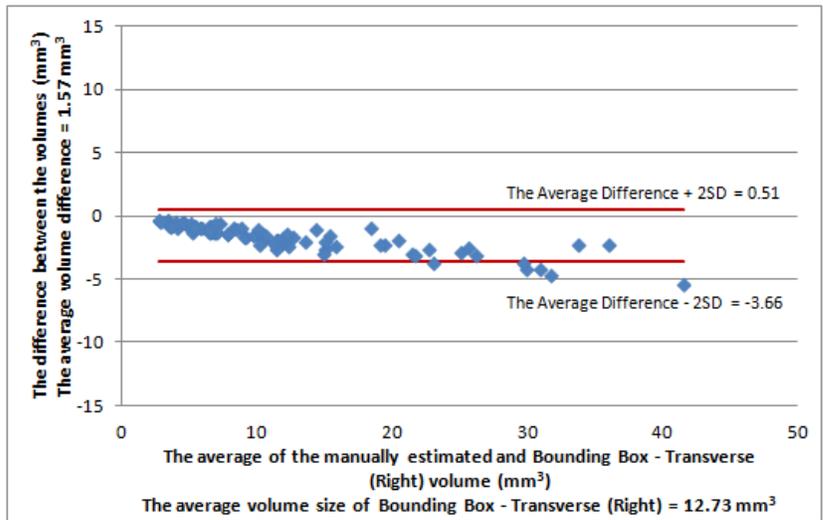


(b) Right Ventricle

Figure 4.14: Bland-Altman plot comparing volumes obtained using manual estimation and the Bounding Box 3-D segmentation algorithm in the Sagittal (SAG) plane

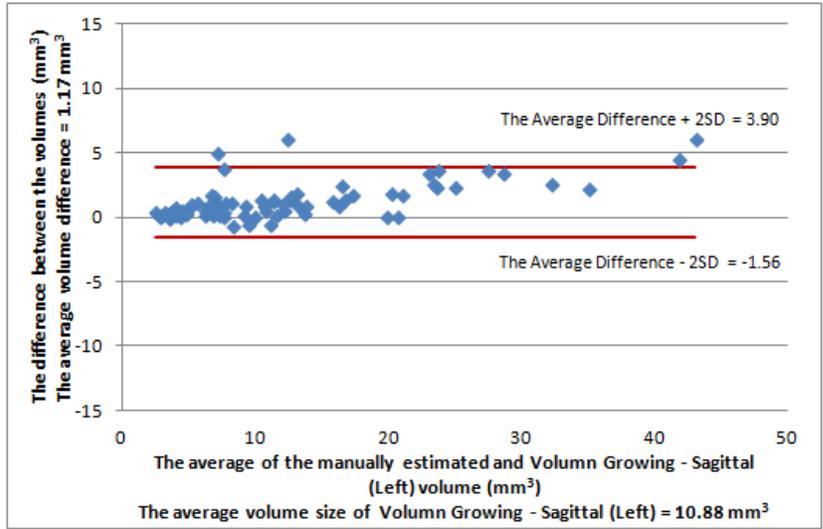


(a) Left Ventricle

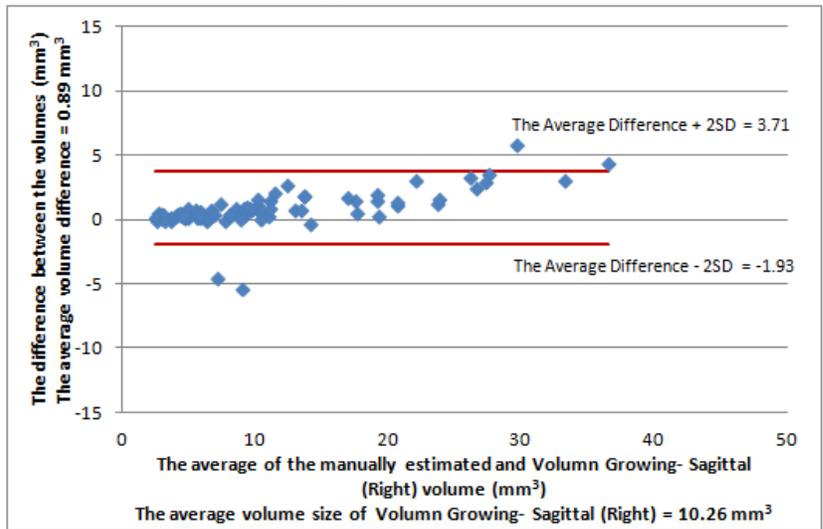


(b) Right Ventricle

Figure 4.15: Bland-Altman plot comparing volumes obtained using manual estimation and the Bounding Box 3-D segmentation algorithm in the Transverse (TRA) plane

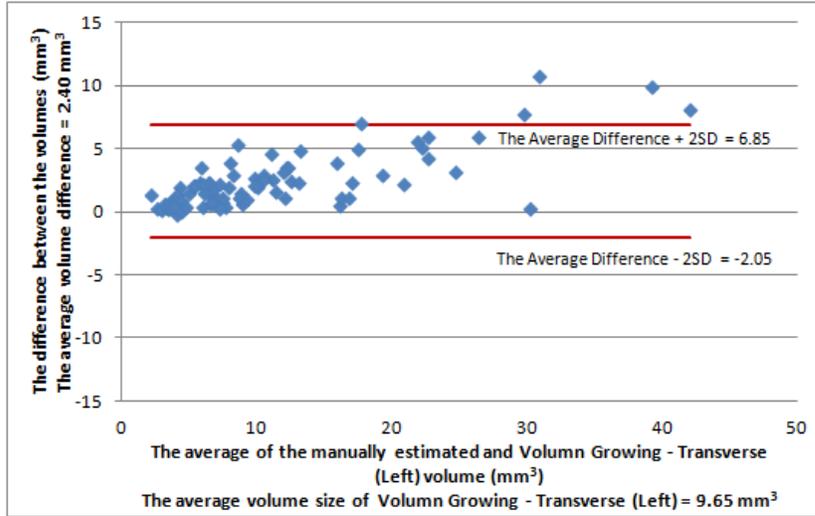


(a) Left Ventricle

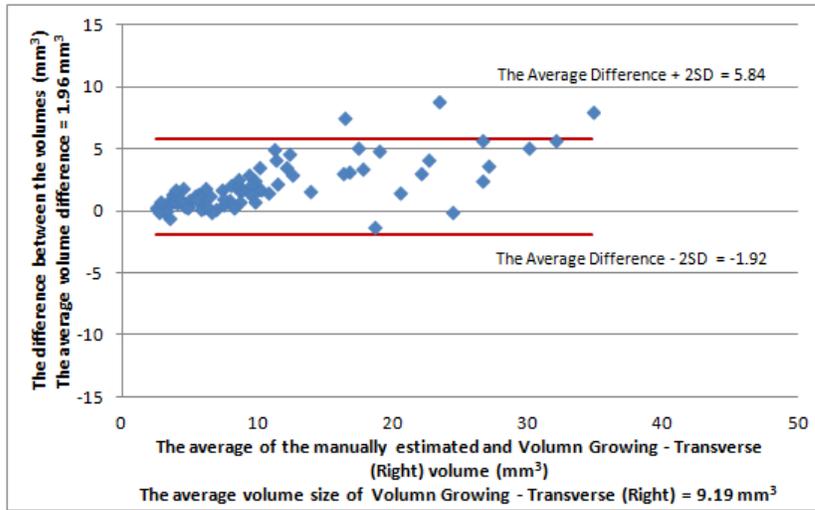


(b) Right Ventricle

Figure 4.16: Bland-Altman plot comparing volumes obtained using manual estimation and the Volume Growing 3-D segmentation algorithm in the Sagittal (SAG) plane



(a) Left Ventricle



(b) Right Ventricle

Figure 4.17: Bland-Altman plot comparing volumes obtained using manual estimation and the Volume Growing 3-D segmentation algorithm in the Transverse (TRA) plane

To give a deeper insight into the results obtained the differences in volume size ( $\text{mm}^3$ ) between the manual and the proposed algorithms are presented using Bland-Altman plots in Figures 4.14 to 4.17. A Bland-Altman plot is a statistical data plotting technique for assessing the agreement or otherwise between two methods of measurement [6]. The mean difference (bias estimate) and the 95 range of agreement (calculated as the mean difference between  $+2\text{SD}$  and  $-2\text{SD}$ ) are represented by the continuous horizontal lines. From the figures it can be clearly seen that the mean difference between the manually estimated volumes and the volumes collected by the Bounding Box technique when used in the SAG plane is the smallest ( $0.49 \text{ mm}^3$  for the left ventricle and  $0.68 \text{ mm}^3$  for the right ventricle) and those collected by the Volume Growing technique when used in the TRA plane is the largest ( $2.40 \text{ mm}^3$  for the left ventricle and  $1.96$

mm<sup>3</sup> for the right ventricle).

Although both of the proposed VOI segmentation algorithms provided the same functionality, there were some different trade-offs between them (other than segmentation effectiveness as demonstrated above). For the Volume Growing-based technique, it required two initial inputs: a starting slice and a “guide point”. The limitation of the algorithm was that the ventricle in the connecting slices might not appear continuously in some cases. Consequently the starting slice and the guide point must be selected with care. In addition, the technique may result in some small errors in cases where the positions of the ventricle between any two connecting slices were significantly different. In this case, the calculated new guide point might not be within the ventricle area in the next slice and consequently the algorithm would terminate prematurely. As a result, some parts of the ventricle might not be collected, which would explain the consistent underestimation (leading to bias) when using the Volume Growing technique. For the Bounding Box-based technique, the box had to be manually defined and thus required some resource. The application of a noise reduction technique was also required in order to remove noise voxels. This had the effect of increasing the runtime of the algorithm. On the positive side the idea behind the technique was simple, easy to implement, and effective.

## 4.4 Conclusion

This chapter has reviewed the necessary preprocessing and segmentation of the 3-D MRI brain scan data used as a focus for the work presented in this thesis. Two 3-D segmentation algorithms were proposed: (i) Volume Growing and (ii) Bounding Box. The evaluation of the two techniques indicated that, although both techniques worked well, the Bounding Box technique tended to be more accurate than the Volume Growing technique (in the context of the size of the volumes identified). Once the VOI have been identified the next stage is to find some mechanism whereby they can be represented in such a way that as little significant information as possible is lost while still permitting the application of data mining techniques (classification with respect to this thesis). In the following three chapters three different representation techniques are considered and evaluated using the lateral ventricle volumes identified using the proposed Bounding Box 3-D segmentation algorithm applied in the SAG plane. Recall that the three representation techniques considered are:

1. Statistical metrics based representation.
2. Point series based representation.
3. Oct-tree based representation.

The first of these, the Statistical metrics based representation, is considered in the next chapter, Chapter 5.

## Chapter 5

# Volume of Interest Image Classification Using Statistical Metrics Based Representation

As noted in the foregoing chapters the application of techniques to classify 3-D image data with respect to some common object that features across a given image set requires the representation of the image objects in question using some appropriate format. In this chapter the Statistical metric based VOI representation is presented, the first of the three representations considered in this thesis, whereby a number of statistical measures are used to represent volumes. The measures are then used to define a N-dimensional feature space, one dimension per statistical measure, from which a feature vector representation can be extracted, one feature vector per volume (ventricle). The feature vectors generated are then used as input to selected classification mechanisms so that classification models can be generated. The aim of the statistical representation, presented in this chapter, is to provide a benchmark approach with which the two alternative representations presented later in this thesis can be compared. The remainder of this chapter is organised as follows. Section 5.1 describes the statistical metrics used. The evaluation of the proposed approach, including discussion, is then presented in Section 5.2. The chapter is summarised in Section 5.3 with some conclusions.

### 5.1 Statistical Metrics Based Image Classification

The usage of statistical techniques, as noted in Chapter 2, is widely reported in the literature. In the context of 2-D, example applications can be found in [50, 52] while in 3-D example applications can be found in [48, 117]. The metrics considered with respect to the work presented in this chapter were:

1. **Axis length ( $l$ ):** The axis length of a given ventricle.
2. **Axis width ( $w$ ):** The axis width of a given ventricle.

3. **Axis depth ( $d$ ):** The axis depth of a given ventricle.
4. **Maximum perimeter in the  $xy$  plane ( $p_{xy}$ ):** The maximum perimeter of a given ventricle in the  $xy$  plane.
5. **Maximum perimeter in the  $yz$  plane ( $p_{yz}$ ):** The maximum perimeter of a given ventricle in the  $yz$  plane.
6. **Maximum perimeter in the  $xz$  plane ( $p_{xz}$ ):** The maximum perimeter of a given ventricle in the  $xz$  plane.
7. **Volume ( $vol$ ):** The volume of a given ventricle (directly related to the number of voxels, note that in the case of the ventricles 1 voxel = 1 mm<sup>3</sup>).
8. **Volume extent ( $v_{ext}$ ):** The value derived by dividing the volume ( $vol$ ) by the size of the minimum bounding cube surrounding the volume, as shown in Equation 5.1.

$$v_{ext} = \frac{Vol}{l \times w \times d} \quad (5.1)$$

Thus we have three length measures, three perimeter measures and two volume measures. These were selected because they seemed to be the most appropriate for use with respect to the VOIBIC concept central to this thesis. The three length measures (axis length, axis width and axis depth) were calculated from the maximum voxel positions on each plane. Note that with respect to the nature of the lateral ventricles two feature vectors were generated in each case, one for the left ventricle and another for the right ventricle. As a result, two sets of feature vectors were produced for each MRI brain scan and used as inputs for classifier generation. The evaluation of the Statistical metrics based approach to VOIBIC is presented in the following section.

## 5.2 Evaluation

This section presents the results obtained with respect to the evaluation of the proposed benchmark VOIBIC Statistical metrics based approach. The image sets used for the evaluation were the Epilepsy and Musician datasets introduced in Chapter 3. The reported metrics are: (i) Accuracy (Accu.), (ii) sensitivity (Sens.) and (iii) specificity (Spec.). Note that Ten-fold Cross Validation (TCV) was used throughout the experimentation. For TCV, the original sample is randomly partitioned into 10 equal sized subsamples. Of the 10 subsamples, a single subsample is retained as the validation data set for testing a generated model, and the remaining 9 subsamples are used as

training data. The cross-validation process is then repeated 10 times (folds), with each of the 10 subsamples used exactly once as the validation data. The 10 results from the folds can then be averaged (or otherwise combined) to produce a single estimation. The number of folds can be adjusted but 10 is most commonly used, hence TCV. All the experiments were conducted using a 2.9 GHz Intel Core i7 with 8GB RAM on OS X (10.9) operating system. The objectives of the evaluation were as follows:

1. To compare the operation of the proposed Statistical metrics based representation in terms of a number of classification models.
2. To determine whether the use of additional meta data (age and gender), would effect the quality of the classification.
3. To compare the relative efficiency of the approach when using meta data and when not using meta data, and when using different classification models.
4. To identify which single statistical features contributed the most to successful classification.

Each of the above objectives is discussed in further detail in the following four subsections, Sub-sections 5.2.1 to 5.2.4. For the first objective four classification methods were considered: (i) artificial Neural Network (NN) [57], (ii) Decision Tree (DT) [96], (iii) Bayesian Network (BN)[46] and (iv) Support Vector Machine (SVM) [25] as provided within the Waikato Environment Knowledge Analysis (WEKA) data mining workbench [54]. The first objective is considered in Section 5.2.1 below, and the second, third and fourth objectives in Subsection 5.2.2, 5.2.3 and 5.2.4 respectively.

### 5.2.1 Usage of Different Classification Models

For the experiments using the four different classification models, and the Epilepsy and Musicians datasets, the results are presented in Tables 5.1 and 5.2 respectively. The results are summarised in bar graph form in Figures 5.1 and 5.2. Note that the ‘NN’, ‘DT’, ‘BN’ and ‘SVM’ acronyms used in the tables indicate the four different classification models considered: artificial Neural Network (NN), Decision Tree (DT), Bayesian Network (BN) and Support Vector Machine (SVM) respectively. Each row in the table shows the classification results obtained for each round of TCV with the average (Avg) and standard Deviation (SD) presented in the bottom two rows of the tables. The best average classification accuracy, sensitivity and specificity are presented using bold font.

From Table 5.1 it can be seen that, with regard to the Epilepsy dataset, the best classification accuracy and specificity, on average, were obtained when using SVM and the best classification sensitivity when using Bayesian Networks. From Table 5.2 it

TCV	NN			DT			BN			SVM		
	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.
1	57.32	62.50	52.38	59.76	65.00	54.76	59.76	65.00	54.76	62.50	52.38	59.76
2	51.22	57.50	45.24	60.75	63.58	57.92	60.19	63.02	57.36	59.76	65.00	54.76
3	62.20	65.00	54.76	58.30	61.13	57.47	57.32	62.50	52.38	58.30	61.13	57.47
4	59.76	65.00	54.76	60.75	63.58	57.92	59.76	65.00	54.76	62.50	52.38	59.76
5	51.22	57.50	45.24	59.76	65.00	54.76	59.76	65.00	54.76	60.75	63.58	57.92
6	57.32	62.50	52.38	58.30	61.13	57.47	60.19	63.02	57.36	59.76	65.00	54.76
7	59.76	65.00	54.76	59.76	65.00	54.76	60.19	63.02	57.36	62.50	52.38	59.76
8	57.32	62.50	52.38	58.30	61.13	57.47	60.75	63.58	57.92	60.75	63.58	57.92
9	59.76	65.00	54.76	60.75	63.58	57.92	57.32	62.50	52.38	59.76	65.00	54.76
10	59.76	65.00	54.76	60.75	63.58	57.92	60.19	63.02	57.36	59.76	65.00	54.76
Avg.	57.56	62.75	52.14	59.71	63.27	56.83	59.54	<b>63.56</b>	55.64	<b>60.63</b>	60.54	<b>57.16</b>
SD	3.67	2.99	3.80	1.07	1.60	1.45	1.21	1.03	2.12	1.45	5.76	2.23

Table 5.1: Classification results of Statistical metrics based classification for the Epilepsy dataset

TCV	NN			DT			BN			SVM		
	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.
1	71.70	73.58	69.81	66.79	72.45	61.13	63.14	70.06	54.29	70.75	71.70	69.81
2	66.98	69.81	64.15	68.30	71.13	65.47	70.75	71.70	69.81	70.75	71.70	69.81
3	66.98	69.81	64.15	70.75	71.70	69.81	70.75	71.70	69.81	68.30	71.13	65.47
4	62.26	66.04	58.49	63.41	70.00	57.14	70.75	71.70	69.81	68.87	71.70	66.04
5	68.87	71.70	66.04	63.14	70.06	54.29	63.14	70.06	54.29	66.98	69.81	64.15
6	68.87	69.81	64.15	68.30	71.13	65.47	66.98	69.81	64.15	68.30	71.13	65.47
7	68.87	62.26	60.38	67.92	68.91	66.04	68.30	71.13	65.47	71.70	73.58	69.81
8	68.87	71.70	66.04	70.75	71.70	69.81	68.30	71.13	65.47	71.70	73.58	69.81
9	68.87	69.81	64.15	68.30	71.13	65.47	66.98	69.81	64.15	70.75	71.7	69.81
10	68.87	70.06	54.29	67.92	68.91	66.04	71.70	73.58	69.81	70.75	71.70	69.81
Avg.	68.87	69.46	63.17	67.56	70.71	64.07	68.08	71.07	64.71	<b>69.89</b>	<b>71.77</b>	<b>68.00</b>
SD	2.43	3.19	4.38	2.57	1.20	5.08	3.09	1.19	5.99	1.64	1.12	2.38

Table 5.2: Classification results of Statistical metrics based classification for the Musician dataset

can be seen that, with regard to the Musician dataset, the best classification accuracy, sensitivity and specificity values, on average, were obtained when using SVM. However, from Figures 5.1 and 5.2, which summarise the results presented un the tables, it can be seen that the classification accuracy among the four classifiers was only marginally different.

### 5.2.2 Usage of Meta Attributes

For the experiments to determine the effect of augmenting the statistical data with meta data experiments were again conducted using the Musician and Epilepsy datasets but only the SVM classification model generator (because experiments reported on in the above sub-section had indicated that SVM classification produced good results).

Table 5.3 shows the classification results obtained using the Statistical metrics based

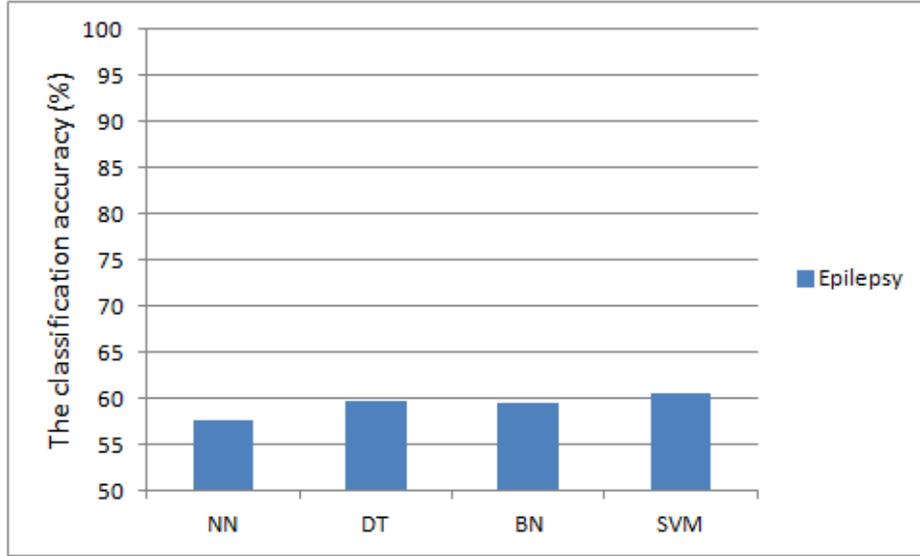


Figure 5.1: Average classification accuracy of Statistical metrics based classification for the Epilepsy dataset

VOIBIC approach for both the Epilepsy and Musician datasets. The same data is summarised in Figure 5.3. With respect to whether it is beneficial to use additional meta data or not, from the table and figure, it can be seen that classification accuracy, sensitivity and specificity, when using augmented data, improved considerably for the Epilepsy dataset and improved slightly for the Musician dataset. It was conjectured that the meta attributes, age and gender, were thus a useful indicator for Epilepsy.

	Augmented						Non Augmented					
	Epilepsy			Musicians			Epilepsy			Musicians		
TCV	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.
1	68.30	71.13	65.47	68.29	75.00	61.90	62.50	52.38	59.76	70.75	71.70	69.81
2	68.30	71.13	65.47	70.73	77.50	64.29	59.76	65.00	54.76	70.75	71.70	69.81
3	64.53	61.70	67.36	70.73	77.50	64.29	58.30	61.13	57.47	68.30	71.13	65.47
4	63.58	67.92	69.25	73.17	80.00	66.67	62.50	52.38	59.76	68.87	71.70	66.04
5	70.19	68.02	67.36	68.29	75.00	61.90	60.75	63.58	57.92	66.98	69.81	64.15
6	69.25	76.79	71.70	68.29	75.00	61.90	59.76	65.00	54.76	68.30	71.13	65.47
7	67.36	75.47	69.25	69.51	75.00	64.29	62.50	52.38	59.76	71.70	73.58	69.81
8	70.19	68.02	67.36	73.17	80.00	66.67	60.75	63.58	57.92	71.70	73.58	69.81
9	70.19	68.02	67.36	68.29	75.00	61.90	59.76	65.00	54.76	70.75	71.70	69.81
10	73.96	76.79	71.13	73.17	80.00	66.67	59.76	65.00	54.76	70.75	71.70	69.81
Avg.	68.59	70.50	68.17	70.36	77.00	64.05	60.63	60.54	57.16	69.89	71.77	68.00
SD	2.99	4.80	2.13	2.16	2.30	2.09	1.45	5.76	2.23	1.64	1.12	2.38

Table 5.3: Classification results of Statistical metrics based classification for Epilepsy and Musician dataset using SVM

It is also interesting to note that, from results presented in Table 5.3 and Figure 5.3, the recorded accuracy, sensitivity and specificity values for the Musician dataset were significantly better than those for the Epilepsy dataset when the meta attributes

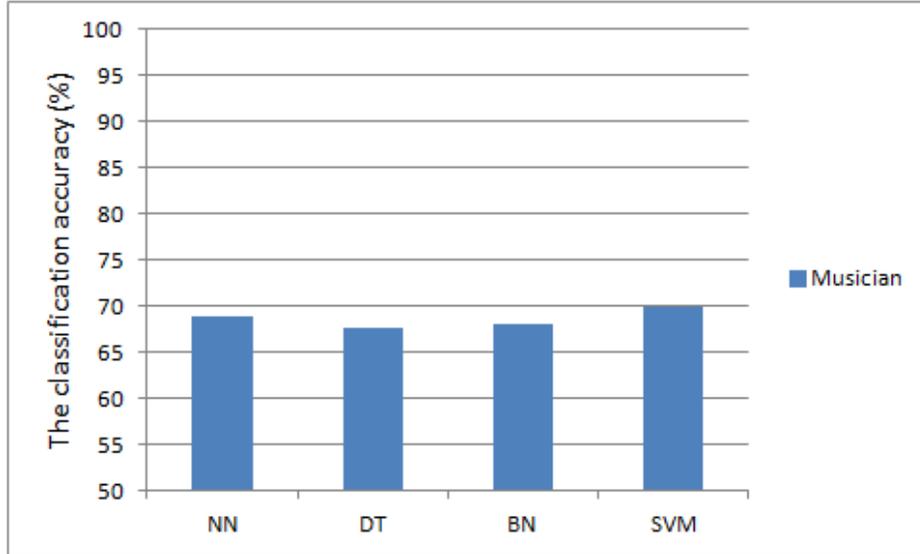


Figure 5.2: Average classification accuracy of Statistical metrics based classification for the Musician dataset

were not used (and only slightly better when the meta attributes were used). It was thus conjectured that the lateral ventricles might be a better indicator of musical ability than epilepsy. El Sayed et al. [40, 37], who published work on the classification of 2-D MRI brain scan data in the context of the corpus callosum, but using the same datasets as used in this thesis, also found this to be true.

### 5.2.3 Efficiency

For the experiments to determine the relevant efficiency of the proposed approach, using augmented and non-augmented data and a number of different classification models, the total run times for each TCV fold was recorded. This thus included classifier model generation time and model testing time. The results (seconds) obtained are presented in Table 5.4. Each row in the table shows the classification time (seconds) for each round of TCV with the average (Avg) and standard deviation (SD) presented in the bottom two rows of the table. Note that the experiments were conducted using a 2.9 GHz Intel Core i7 with 8GB RAM on OS X (10.9) operating system.

From the table it can be seen that, on average, the recorded run time using SVM classifier generation was the fastest for the Epilepsy Augmented, the Epilepsy Non-Augmented and the Musician Non-Augmented datasets; while the run time using Bayesian Network was the fastest for the Musician Augmented dataset. Using Decision Trees produced the worst run time results for the Epilepsy Non-Augmented and the Musician Non-Augmented dataset; while using Neural Networks produced the worst results with respect to the Epilepsy Augmented and the Musician Augmented datasets. It can also be seen that, on average, the run times with respect to the Musi-

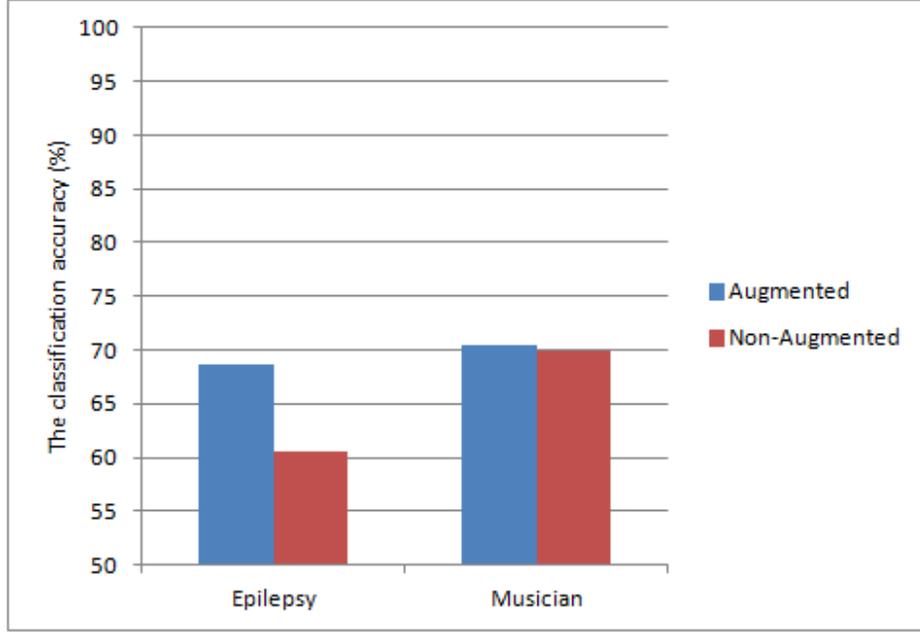


Figure 5.3: Average classification accuracy of Statistical metrics based classification using SVM

	Non-Augmented								Augmented							
	Epilepsy				Musician				Epilepsy				Musician			
TCV	SVM	DT	BN	NN	SVM	DT	BN	NN	SVM	DT	BN	NN	SVM	DT	BN	NN
1	12.30	14.30	12.40	14.40	8.40	10.20	9.00	10.20	13.40	16.00	14.10	16.50	10.00	12.00	9.50	14.50
2	12.20	15.10	13.10	14.30	8.00	9.30	9.10	9.50	14.00	14.50	14.10	17.00	10.20	13.20	11.00	15.00
3	12.50	15.20	14.20	15.40	9.00	10.20	8.40	9.50	13.00	15.50	14.20	15.40	10.20	12.30	10.20	15.10
4	11.20	16.10	13.00	13.50	9.10	9.50	9.20	9.50	12.50	16.00	14.50	15.50	10.30	12.50	9.50	15.10
5	12.10	15.50	13.40	14.00	9.30	10.00	9.20	10.00	12.40	15.00	14.50	16.00	10.20	12.50	10.00	15.20
6	11.30	14.50	12.10	13.00	9.00	10.10	9.30	10.10	12.30	15.10	15.40	16.40	10.40	12.50	10.20	15.20
7	12.20	15.30	12.50	13.30	9.10	10.20	9.00	10.30	12.40	15.30	15.00	16.40	10.30	13.30	10.20	15.20
8	12.30	16.00	12.50	13.40	8.20	10.50	9.10	10.30	13.30	15.30	14.30	17.00	10.40	13.00	10.30	15.20
9	12.40	15.40	12.50	14.30	8.50	9.50	9.50	9.50	13.50	15.40	14.50	17.00	10.40	12.50	10.30	15.00
10	12.50	15.30	13.10	14.00	9.00	10.00	9.00	10.00	13.40	15.40	15.00	16.50	11.00	13.00	10.20	15.50
Avg.	12.10	15.27	12.88	13.96	8.76	9.95	9.08	9.89	13.02	15.35	14.56	16.37	10.34	12.68	10.14	15.10
SD	0.47	0.56	0.61	0.70	0.45	0.39	0.29	0.35	0.59	0.45	0.44	0.58	0.26	0.42	0.43	0.25

Table 5.4: Total run time (seconds) for classification (model generation and testing)

Musician datasets was faster than those for the Epilepsy datasets. This was clearly because the total number of MRI brain scans in the Epilepsy dataset (210 brain scans) was more than those in the Musician dataset (160 brain scans). The average run time with respect to each classifier generation model is given in Figure 5.4. Note that ‘Epilepsy’, ‘Epilepsy+’, ‘Musician’ and ‘Musician+’ indicate the Epilepsy dataset, the Epilepsy dataset with augmented data, the Musician dataset and the Musician dataset with augmented data respectively.

The results presented in Table 5.4 did not include the time required to generate the representation. The run time required to generate the statistical metrics based representation was 157 seconds for the Epilepsy dataset 112 seconds for the Musician dataset. Thus, as expected, the generation time required for the Musician dataset was

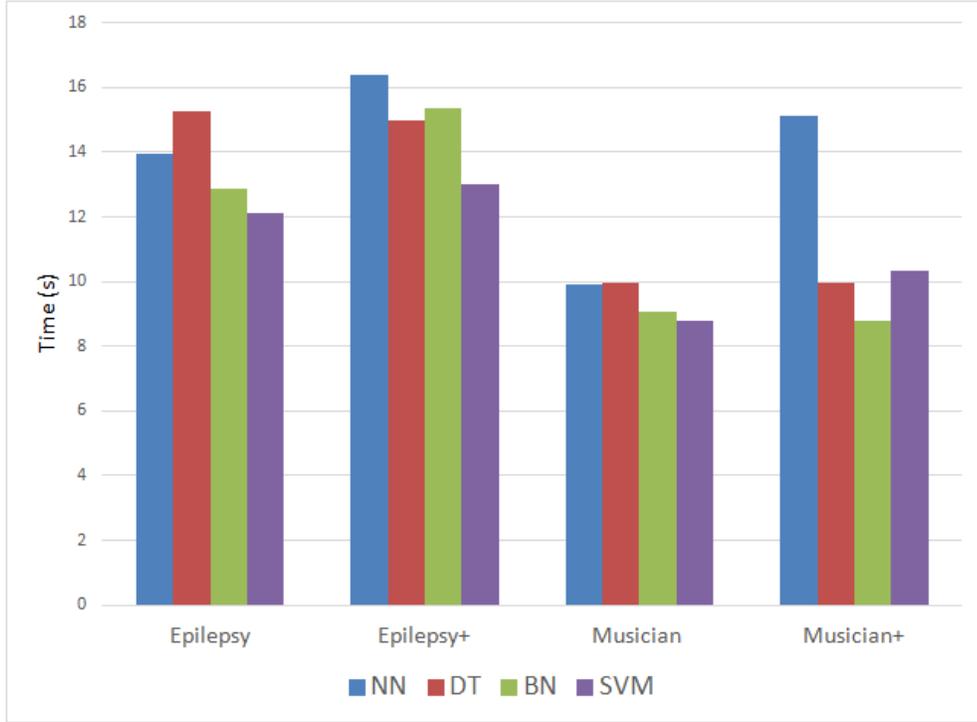


Figure 5.4: Average run time complexity of Statistical metrics based image classification

less than that required for the Epilepsy dataset (because the total number of MRI brain scans was less than the total number of Epilepsy brain scans). The processing time per brain scan volume for each data set was 0.7 seconds.

For completeness the total time complexity with regard to the Statistical metrics based classification (statistical metric representation generation time plus classification time) is shown in Table 5.5 and Figure 5.5. From the table and figure it can be seen that the total run time complexity for the Musician dataset was the best and that for the Epilepsy dataset with augmented data was the worst. Note that the run times presented in the table and figure were generated using the best performing classifier, namely SVM.

Dataset	Representation time (s)	Classification time (s)	Total time (s)
Epilepsy	157	12.10	169.10
Epilepsy+	157	13.02	170.02
Musician	112	8.76	120.86
Musician+	112	10.14	122.14

Table 5.5: The total run time using Statistical metrics based classification (representation generation time plus classification time)

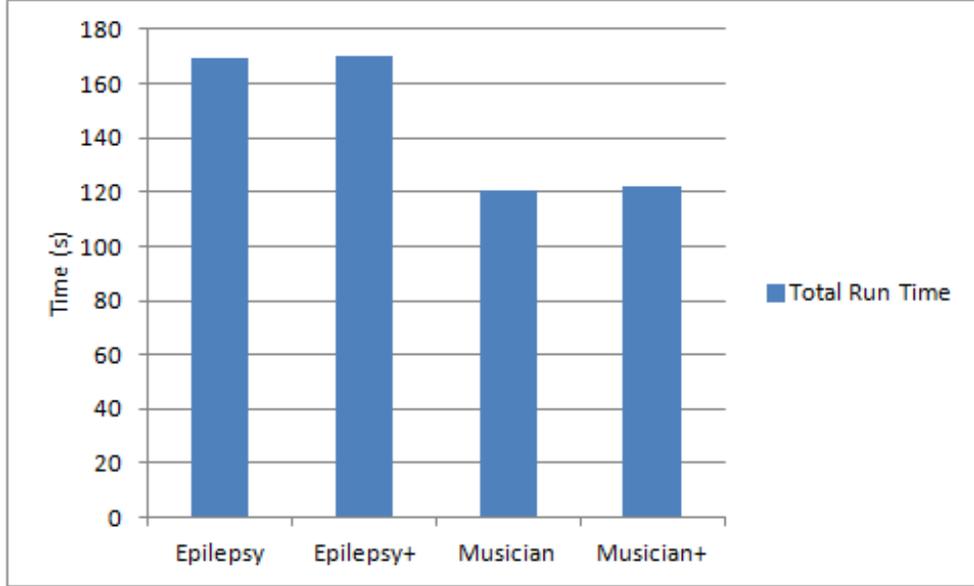


Figure 5.5: The total run time using Statistical metrics based classification (representation generation time plus classification time)

#### 5.2.4 Comparison of Attributes

For the experiments to determine which single attribute contributed the most to the classification of the left and right ventricles each attribute was considered in isolation. Each attribute was therefore used to generate a 1-D feature vector to represent the identified VOI. TCV was again adopted. Only results obtained using the SVM classification model are reported here because earlier experiments, reported above, had demonstrated that SVM produced the best results. The results are presented in Table 5.6 (best results highlighted in bold font). From the table it can be seen that “volume extent” consistently provided the best results; while axis length and depth consistently provided the worst results.

Attribute	Epilepsy			Musician		
	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.
Axis length ( $l$ )	51.22	57.50	45.24	44.34	46.23	42.45
Axis width ( $w$ )	53.77	58.49	49.06	51.38	55.15	47.60
Axis depth ( $d$ )	48.78	52.50	45.24	48.21	52.92	45.49
Maximum perimeter on $xy$ plane ( $p_{xy}$ )	54.72	58.49	50.94	56.60	59.43	53.77
Maximum perimeter on $yz$ plane ( $p_{yz}$ )	56.60	62.26	54.72	54.62	59.81	49.43
Maximum perimeter on $xz$ plane ( $p_{xz}$ )	51.98	54.72	49.06	50.38	54.15	46.60
Volume ( $vol$ )	50.31	49.06	50.94	55.09	59.81	50.38
Volume extent ( $v_{ext}$ )	<b>64.15</b>	<b>64.15</b>	<b>64.15</b>	<b>59.76</b>	<b>65.00</b>	<b>54.76</b>

Table 5.6: Classification results using Statistical metrics based classification, single attribute feature vectors and SVM

### 5.3 Conclusion

In this chapter a statistical metrics based approach to VOIBIC has been described. The Statistical metrics based classification approach used the geometric attributes of the lateral ventricles to generate feature vectors. Four classification mechanisms, SVM, Decision Tree, Artificial Neural Network and Bayesian Network were considered with which to evaluate the proposed approach. The recorded results indicated that a best accuracy of 70.36% was produced. The results also demonstrated that SVM classification, on average, was the best classification model to adopt in terms of effectiveness when augmenting the data with the meta attributes, “age” and “gender”. Considering the statistical metrics individually, volume extent provided the best results, while ventricle width and depth produced the worst results. The main purpose of the statistical VOIBIC approach presented in this chapter was to establish a benchmark approach. In the following Chapter an alternative approach to classifying 3-D MRI scans using a Point series based image representation technique, that outperforms the proposed Statistical metrics based techniques presented in this chapter, will be presented.

## Chapter 6

# Volume of Interest Image Classification Using Point Series Representation

In this chapter, the proposed approach to VOIBIC founded on the concept of a point series representation is presented. A point series in this case is a series (sequence) of points used to represent an object in terms of (say) its perimeter or circumference. With respect to the work in this chapter the point series represent the boundary features of the VOI. Two techniques for generating the desired point series are considered: (i) Disc based and (ii) Spoke based. Regardless of which technique is used the resulting point series are represented as curves which can be represented as graphs where the x-axis is the order (or sequence) of the points and the y-axis is the distance values associated with the points. These can be translated into a feature vector representation to which a number of classification model generation mechanisms can be applied, as in the case of the statistical metrics based representation discussed in the previous chapter, or alternatively they can be used directly using a K-Nearest Neighbour (KNN) approach. For the first, a feature space representation is required. It is proposed in this chapter that this be generated using Hough signature extraction [62]. For the KNN mechanism a similarity measure is required. The simplest similarity measure that can be used is the Euclidean distance between a labelled comparator series and a previously unseen series. However, this requires that both series are of the same length. Instead it is proposed that the “warping path” distance, generated using Dynamic Time Warping (DTW) [11], is the most appropriate for this purpose.

The chapter is organised as follows. Section 6.1 describes the two techniques for generating the point series from a given VOI (ventricle). This is followed in Section 6.2 with a description of the suggested classification processes to be adopted. The evaluation of the proposed techniques, including a discussion of the results obtained, is then presented in Section 6.3. Finally, the chapter is summarised in Section 6.4.

## 6.1 Point Series Image Representation

As noted in the introduction to the chapter, the focus of the work presented is the representation of VOIs (lateral ventricles in particular), in terms of their boundary, using point series (curves). The conjecture was that this would serve to capture the nature of the shape of the ventricles in a manner commensurate with prediction/classification. The challenges of generating such point series are:

1. **Identification/selection of the points to be included in the point series.**  
It is unclear whether it is best to include all boundary voxels in the desired point series or some subset of the voxels. The argument for using all voxels is that all the data is being used (nothing is being left out). The argument for using only a subset of the voxels is that it is neither necessary or practical (from an efficiency perspective) to use all the data. Thus we need to consider mechanisms to achieve both.
2. **Imposing an ordering on the identified/selected points.** In 2-D it is straightforward to impose an ordering given a collection of boundary pixels, we simply select a point as the start point and then proceed in either a clockwise or anti-clockwise direction until we return to the start point (as in the case, for example, of chain coding [44]). However, in 3-D what this ordering might be is less obvious.

As also noted in the introduction to this chapter two techniques for generating the desired point series are proposed: (i) Disc based and (ii) Spoke based. The first uses all the boundary voxels while the second uses a representative subset of the complete set of boundary voxels. The input in both cases is a binary-valued 3-D image comprising voxels labelled as being either black (belonging to the VOI) or white (not part of the VOI). The output in both cases is a point series describing the VOI boundary. Both the Disc based and Spoke based techniques operate with reference to a *primary axis* selected from one of the three available axis: (i) Sagittal ( $xy$ ), (ii) Coronal ( $yz$ ) and (iii) Transverse ( $xz$ ). Consequently for each image three point series can be generated for each ventricle. Thus, in total six curves are generated for each MRI image, three describing the left ventricle and three describing the right ventricle. The distinction between the two techniques is how the curves are generated. Each technique is described in further detail in the following two subsections.

### 6.1.1 Disc-based Representation Technique

The Disc-based representation is founded on the idea of generating a point series by considering a sequence of slices, slice by slice (along a primary axis), and collecting point information from the boundary where each slice and the volume of interest intersect.

The intersection is usually described by a circular shape, as illustrated in Figure 6.1 hence the technique is referred to as the “Disc” based technique. In this manner a sequence of disc boundaries is generated and concatenated together to form a single point series. In more detail, the Disc-based point series generation process is as follows:

1. Find the geometric centroid of the ventricle under consideration.
2. Starting at one end of the VOI (ventricle), with respect to the selected primary axis, define the first “disc”  $d$ .
3. Calculate the distances from the identified centroid to the boundary points (voxels) describing the edge of  $d$ , obtained from the intersection of the disc and the boundary of the VOI, and store the distances.
4. Define a follow on slice by moving one voxel along the selected axis and repeating (3) until the entire VOI has been defined. Note that on start up there will be two follow on slices one on either side of the centroid.
5. Use the collected distances to define a curve (point series) with distance along the Y-axis and the sequential point numbers along the X-axis.

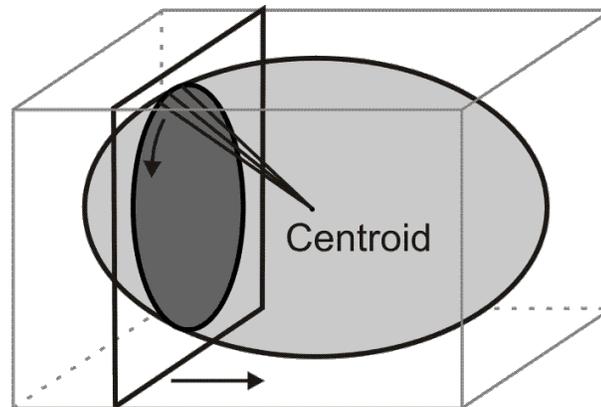


Figure 6.1: Representation of a ventricle (VOI) as a point series using the Disc-based technique

An example of a curve generated using the Disc-based technique is given in Figure 6.2.

### 6.1.2 Spoke-based Representation Technique

The Spoke-based representation technique is illustrated in Figure 6.3. The technique involves measuring the distance from the geometric centroid of the VOI to points on the boundary, in a given plane. As such it is essentially a 2-D technique. The effect is that of a sequence of spokes of different length radiating from the centroid, hence the name of the technique. The generation process is as follows:

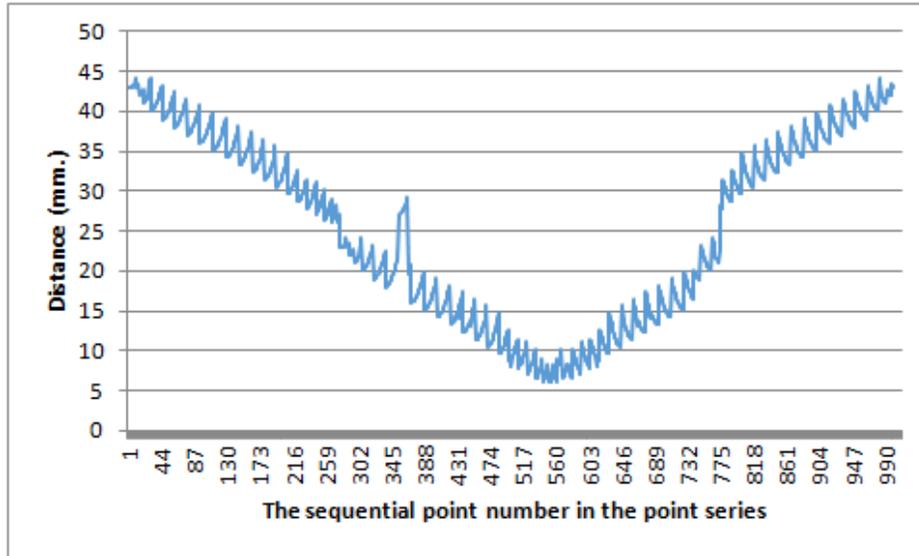


Figure 6.2: An example of a point series curve generated using the Disc-based technique

1. Finding the centroid of the ventricle.
2. Generate a spoke, in some pre-define plane, radiating from the centroid to the edge of the VOI (ventricle) and measure and record its length.
3. Repeat (2) with a new spoke generated by rotating the previous spoke by an angle of  $\theta^\circ$ . Continue in this manner until all  $360^\circ$  have been covered.
4. Use the collected distances to define a curve (point series) with spoke-distance along the Y-axis and the sequential point numbers along the X-axis.

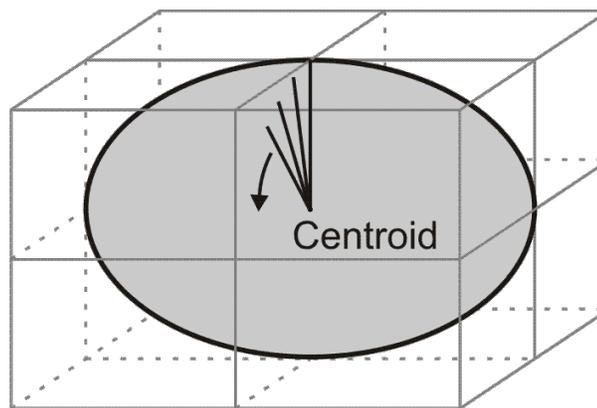


Figure 6.3: Representation of a ventricle (VOI) as a point series using the Spoke-based technique

An example of a curve generated using the Spoke-based technique is given in Figure 6.4.

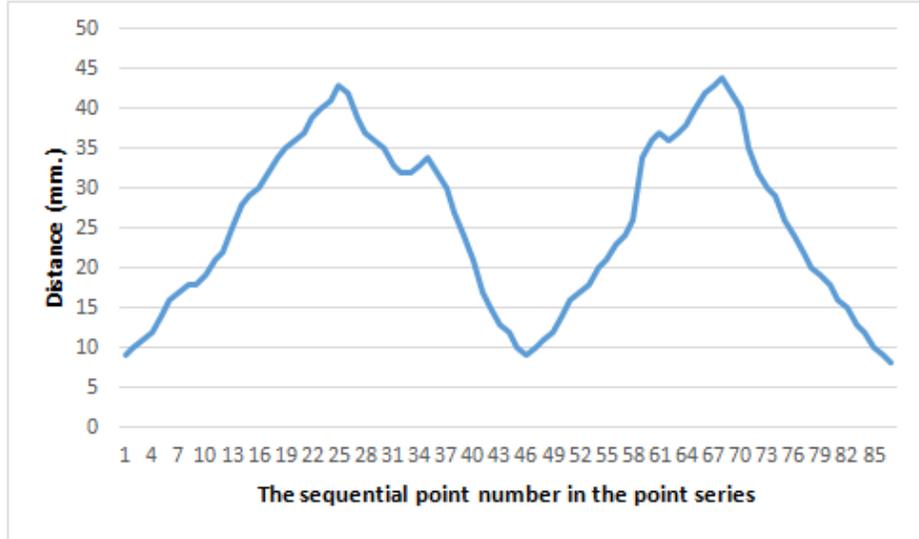


Figure 6.4: An example of a point series curve generated using the Spoke-based technique

## 6.2 Framework for Image Classification Based on Point Series Representation

The point series based image classification approach starts with the identification of the VOI as described in Chapter 4. Once the VOI have been identified the next step is to generate point series describing the VOI using one or other of the techniques presented in Section 6.1. After that the classification process is conducted using either: (i) Feature space classification or (ii) KNN. Each is explained in further detail in the following two sub-sections, Sub-sections 6.2.1 and 6.2.2.

### 6.2.1 Feature Space Classification

Using feature space classification the curves generated, as described above, are processed to create a feature space from which feature vectors can be generated (one per image). More specifically a signature based approach is proposed for generating the model, founded on Hough signature extraction [62]. The generation of shape signatures using the concept of the Hough Transform, as used in this thesis, was first proposed by Paul Hough in 1962 [62]. The idea was widely adopted and became the basis for a great number of image analysis techniques with respect to many application domains. The Hough transform is mainly used for parametric shapes in images. The key idea is to transform a shape pattern into a parameter space where the shape can be represented in a spatially compact way. With regard to the work presented in this thesis the Straight Line Hough Transform (SLHT) was used, the first, and probably most used, of the parameter-based transformations derived from the Hough concept. A signature in this context is a set of feature values that can be used to describe some entity, a

curve in our case. The feature values encompassed by a set of signatures thus describe a feature space.

The process of generating the desired signatures commences by transforming the curves (generated using one or other of the techniques described in Section 6.1.1 and 6.1.2) into a parameter space (accumulator matrix)  $A$ , comprised of  $M$  rows and  $N$  columns where  $M$  is the maximum recorded (Disc or Spoke) distance and  $N$  is the total number of points in the point series.  $\rho$  and  $\theta$  are the parameters of the corresponding curve in the parameter space  $(\rho, \theta)$ , where  $\theta$  varies from the minimum to the maximum value, giving the corresponding  $\rho$  value. Each row corresponds to one value of  $\rho$ , and each column corresponds to one value of  $\theta$ . The process for generating a signature from an accumulator matrix, as first proposed in [130], is described as follows:

1. Create accumulator matrix  $A$  comprised of  $M$  rows and  $N$  columns.
2. Put “0” into all the cells in the accumulator matrix  $A$ .
3. Set  $\rho = “1”$ .
4. If  $\rho$  greater than  $N$ , go to (7).
5. Put “1” into the cell  $(\rho, \theta)$ , where  $\theta$  is the distance value corresponding to  $\rho$  ( $\rho$  is the order of point series).
6. Increment  $\rho$  and go back to (4).
7. Calculate a preliminary feature vector from the accumulator matrix  $A$ :

$$F_j = \sum_{i=1}^M A_{ij}^2, j = 1..N$$

An example of the generation of accumulator matrix can be seen in Figure 6.5 and Table 6.1

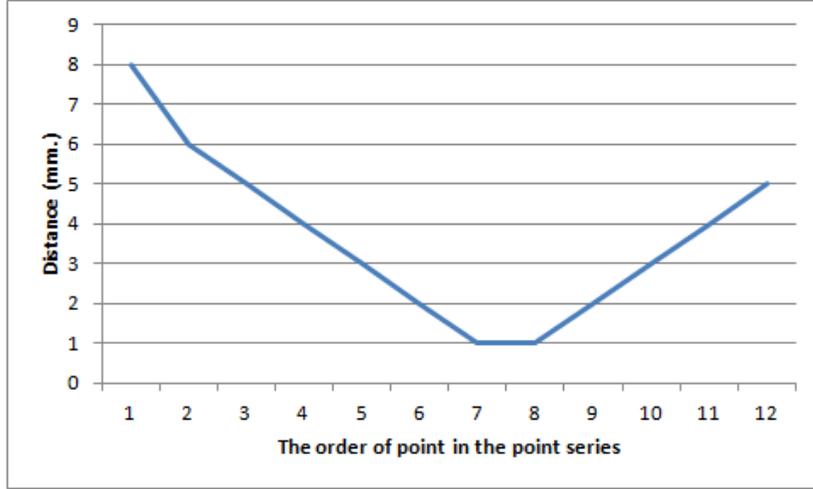


Figure 6.5: Example of a curve  $Z$

8	1	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0
6	0	1	0	0	0	0	0	0	0	0	0	0
5	0	0	1	0	0	0	0	0	0	0	0	1
4	0	0	0	1	0	0	0	0	0	0	1	0
3	0	0	0	0	1	0	0	0	0	1	0	0
2	0	0	0	0	0	1	0	0	1	0	0	0
1	0	0	0	0	0	0	1	1	0	0	0	0
	1	2	3	4	5	6	7	8	9	10	11	12

Table 6.1: Example of the accumulator matrix generated from the curve  $Z$  in Figure 6.5

8. Calculate the vector mean:

$$\mu = \frac{1}{N} \sum_{j=1}^N F_j$$

9. Normalise the feature vector:

$$FV_j(\theta) = \frac{F_j}{\mu}, j = 1..N$$

In this process, the feature vector is calculated from the accumulator matrix  $A$  in step (7) and is normalised according to its mean in steps (8) and (9). The extracted feature vector describing the curve, which reflects the shape of the VOI, can finally be used as a signature. The extracted signatures are then stored together with an associated class label to which any established classifier generation mechanism can be applied.

## 6.2.2 KNN Classification

Using KNN the set of labelled curves defining a VOI can be used directly, without transition into an intermediate representation as in the case of the feature space classification technique describe above, to classify unseen data as described in [33]. This requires some kind of measure to determine the similarity between curves in the “curve base” and a new unseen curve. The DTW algorithm is well-known in many areas. It is a well tried and tested technique first described in the 1960s and extensively researched in 1970s for application with respect to speech recognition systems. It has more recently been used in the context of handwriting recognition systems [87], online signature matching systems [36] and protein sequence alignment [127]. As noted earlier warping distance as generated using the DTW algorithm [11] was used with respect to this thesis.

DTW operates in two stages: (i) matrix generation and (ii) warping path calculation. Algorithm 1 gives the matrix generation process. The input is two curves  $X$  and  $Y$ , the first of length  $m$  and the second of length  $n$ . We commence by stepping through curve  $X$  and for each element  $x_i$  in curve  $x$  stepping through curve  $Y$ . For each element  $y_j$  in  $Y$  we find the difference with  $x_i$  and store this in Matrix  $A$  a location  $[i, j]$ . The input to Algorithm 2 is the matrix  $A$  from Algorithm 1. The output is the warping path distance.

---

**Algorithm 1** The algorithm to generate time warping matrix

---

```
1: procedure MATRIXGENERATION
2:   Input:  $X = \{x_1, x_2, \dots, x_m\}$  and  $Y = \{y_1, y_2, \dots, y_n\}$ 
3:   Output: Matrix  $A$ 
4:   For  $i := 1$  to  $m$  do
5:     For  $j := 1$  to  $n$  do
6:        $A[i, j] \leftarrow x_i \sim y_j$             $\triangleright A[i, j]$  records the difference between  $x_i$  and  $y_j$ 
7:     EndFor
8:   EndFor
```

---

---

**Algorithm 2** Algorithm to generate minimum warping path from Matrix A

---

```
1: procedure DTW
2:     Input: Matrix  $A$ 
3:     Output:  $wp$  (Warping Path distance)
4:      $x \leftarrow 0$ 
5:      $y \leftarrow 0$ 
6: While ( $y \neq m$  and  $x \neq n$  do)
7:     if ( $x == m$ ) then
8:          $y \leftarrow y + 1$ 
9:     else if ( $y == n$ ) then
10:         $wp = wp + A[x, y]$ 
11:    else
12:         $dist1 = A[x, y + 1]$ 
13:         $dist2 = A[x + 1, y + 1]$ 
14:         $dist3 = A[x + 1, y]$ 
15:        if ( $dist1 < dist2$ ) then ▷ Distance 1 less than distance 2
16:            if ( $dist1 < dist3$ ) then ▷ Distance 1 less than distances 2 and 3
17:                 $wp \leftarrow wp + dist1$ 
18:                 $y \leftarrow y + 1$ 
19:            else ▷ Distance 1 less than distance 2 but not less than distance 3
20:                 $wp \leftarrow wp + dist3$ 
21:                 $x \leftarrow x + 1$ 
22:        else if ( $dist2 < dist3$ ) then ▷ Distance 2 less than distances 1 and 3
23:             $wp = wp + dist2$ 
24:             $x \leftarrow x + 1$ 
25:             $y \leftarrow y + 1$ 
26:        else ▷ Distance 2 less than distance 1 but not less than distance 3
27:             $wp \leftarrow wp + dist3$ 
28:             $x \leftarrow x + 1$ 
29: EndWhile
```

---

Thus, given two curves,  $X$  with length  $m$  and  $Y$  with length  $n$ , a matrix  $A$  is constructed with  $m$  rows and  $n$  columns. Each element  $(i, j)$  within matrix  $A$  describes the distance between point  $i$  on curve  $X$  and the point  $j$  on curve  $Y$ . The goal is to find the “warping path” through this matrix describing the shortest distance from  $(0, 0)$  to  $(m, n)$ . For instance, given two point series (curves)  $X = [1, 1, 2, 3, 2, 0]$  and  $Y = [0, 1, 1, 2, 3, 2, 1]$ , the matrix  $A$  is constructed as shown in Table 6.2. In the matrix the  $x$ -axis represents curve  $X$  and the  $y$ -axis represents curve  $Y$ . The sequence, marked in red colour, close to the diagonal, indicates the “warping path”. Using Algorithm 2 the algorithm starts from  $A[0, 0]$  and finds the neighbour which has the minimum distance value. Note that the algorithm can move forward only, which means that from a point  $(x, y)$ , the path can only move to right  $(x + 1, y)$ , upwards  $(x, y + 1)$  or diagonal  $(x + 1, y + 1)$ . The algorithm then stops when the  $A[m, n]$  is reached.

According to the Table 6.2, the number of row  $M$  is the length of curve  $X$  and the number of column  $N$  is the length of curve  $Y$ . To find the warping path the

<b>1</b>	0	0	1	2	1	<b>1</b>
<b>2</b>	1	1	0	1	<b>0</b>	2
<b>3</b>	2	2	1	<b>0</b>	1	3
<b>2</b>	1	1	<b>0</b>	1	0	2
<b>1</b>	0	<b>0</b>	1	2	1	1
<b>1</b>	<b>0</b>	0	1	2	1	1
0	<b>1</b>	1	2	3	2	0
	<b>1</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>2</b>	<b>0</b>

Table 6.2: Example of Matrix  $A$  considering distance between curve  $X$  and  $Y$

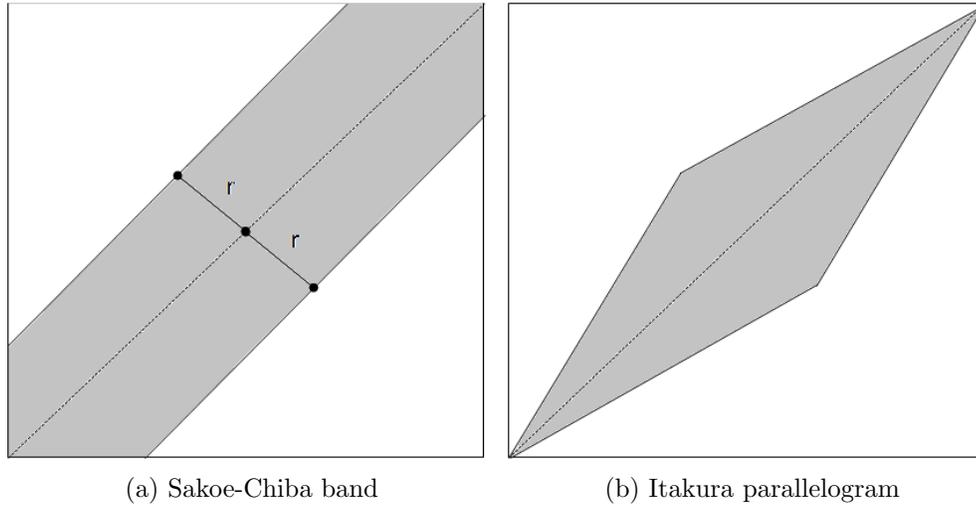


Figure 6.6: Global warping path constraints

algorithm computes no more than  $M$  multiplied by  $N$  times. Hence,  $O(MN)$  is the computational cost of the application of DTW. In order to improve the efficiency, global constraints may be introduced whereby we ignore matrix locations away from the main diagonal. Two well known global constraints are the “Sakoe-Chiba band” [106] and “Itakura parallelogram” [67]. As shown in Figure 6.6, alignments of points can be selected only from the respective shaded regions. The Sakoe-Chiba band runs along the main diagonal and has a fixed width  $R$  such that  $j - R \leq i \leq j + R$  for the indices of the warping path  $w_k(i, j)$  (see Figure 6.6a). While the Itakura parallelogram describes a region that constraints the slope of the warping path. More specifically, for a fixed  $S > 1$ , the Itakura parallelogram consists of all points that are traversed by some warping path having a slope between  $1/S$  and  $S$  (see Figure 6.6b). There are several reasons for using global constraints, one of which is that they slightly improve the performance of the efficiency of the DTW distance calculation. However, the most important reason is to prevent pathological warpings, where a relatively small section of one time series maps onto a relatively large section of another. With respect to the work presented in this thesis the “Sakoe-Chiba” band was adopted.

### 6.3 Evaluation

A number of experiments were conducted, designed to compare the operation of both the Disc-based and Spoke-based point series generation techniques coupled with both Feature space and KNN classification. The objectives of the evaluation were as follows:

1. To compare the operation of the Disc-based point series generation with the Spoke-based point series generation (recall that the Disc-based approach used all the boundary voxels while the Spoke-based approach used a representative subset).
2. To identify the most appropriate value for  $\theta$  in the context of the Spoke-based point series generation approach (recall that the greater the value of  $\theta$  the fewer the number of points in our point series).
3. To compare the usage of Feature space classification against KNN classification in the context of the proposed point series representation.
4. To determine whether the use of additional meta data, age and gender, would effect the quality of the classification.

The image sets used for the evaluation were the Epilepsy and Musician datasets introduced in Chapter 3. Ten-fold Cross Validation (TCV) was used throughout the experimentation. All the experiments were conducted using a 2.9 GHz Intel Core i7 with 8GB RAM on OS X (10.9) operating system. The overall effectiveness results are presented in Tables 6.3 to 6.6. These are discussed in Section 6.3.1, 6.3.2, 6.3.3 and 6.3.4, with respect to the above objectives. Tables 6.3 and 6.4 show the classification results obtained using the Feature space classification approach for the Epilepsy and Musician datasets respectively, while Tables 6.5 and 6.6 show the results obtained using KNN classification approach for the Epilepsy and Musician datasets respectively. Note that ‘Accu.’, ‘Sens.’ and ‘Spec.’ refer to accuracy, sensitivity and specificity respectively.

Technique	Accu.	Sens.	Spec.
Disc	59.43	58.49	<b>60.38</b>
Spoke ( $x = 1^\circ$ )	62.20	67.50	57.14
Spoke ( $x = 2^\circ$ )	<b>64.63</b>	<b>70.00</b>	59.52
Spoke ( $x = 3^\circ$ )	61.32	62.26	<b>60.38</b>
Spoke ( $x = 4^\circ$ )	58.49	62.26	57.42
Average	61.21	64.10	58.97

Table 6.3: Classification results for the Epilepsy dataset using Feature space classification

Technique	Accu.	Sens.	Spec.
Disc	73.58	75.47	72.64
Spoke ( $x = 1^\circ$ )	78.62	83.02	<b>76.42</b>
Spoke ( $x = 2^\circ$ )	<b>79.25</b>	<b>84.91</b>	<b>76.42</b>
Spoke ( $x = 3^\circ$ )	74.21	75.47	73.58
Spoke ( $x = 4^\circ$ )	68.55	69.81	67.92
Average	74.84	77.74	73.40

Table 6.4: Classification results for the Musician dataset using Feature space classification

Technique	Accu.	Sens.	Spec.
Disc	62.20	67.50	57.14
Spoke ( $x = 1^\circ$ )	64.15	66.04	62.26
Spoke ( $x = 2^\circ$ )	<b>69.81</b>	71.70	<b>67.92</b>
Spoke ( $x = 3^\circ$ )	68.87	<b>75.47</b>	62.26
Spoke ( $x = 4^\circ$ )	60.98	67.50	57.14
Average	65.20	69.64	61.34

Table 6.5: Classification results for the Epilepsy dataset using KNN classification

Technique	Accu.	Sens.	Spec.
Disc	77.36	81.13	75.47
Spoke ( $x = 1^\circ$ )	81.13	86.79	78.30
Spoke ( $x = 2^\circ$ )	<b>82.39</b>	<b>88.68</b>	<b>79.25</b>
Spoke ( $x = 3^\circ$ )	76.10	79.25	74.53
Spoke ( $x = 4^\circ$ )	69.81	71.70	68.87
Average	77.36	81.51	75.28

Table 6.6: Classification results for the Musician dataset using KNN classification

### 6.3.1 Disc-Based versus Spoke-Based

This section gives a discussion of the classification results obtained using Disc-based and Spoked-based point series generation. From Tables 6.3 and 6.4 it can be seen that, with respect to Feature space classification, the best classification accuracy and sensitivity were obtained using the Spoke-based image representation with a Spoke spacing of  $\theta = 2^\circ$ . The best classification specificity was obtained using both the Disc-based representation and the Spoke based representation with a spacing of: (i)  $\theta = 3^\circ$  for the Epilepsy dataset and (ii)  $\theta = 1^\circ$  and  $\theta = 2^\circ$  for the Musicians dataset. With respect to the KNN classification technique, from Tables 6.5 and 6.6, it can be seen that for the Epilepsy dataset, the best classification accuracy and specificity were obtained using the Spoke-based approach with a spacing of  $\theta = 2^\circ$ , while the best

sensitivity was obtained using the Spoke-based approach and a spacing of  $\theta = 3^\circ$ . For the Musician dataset, All the best classification accuracy, sensitivity and specificity values were obtained using a Spoke spacing of  $\theta = 2^\circ$ .

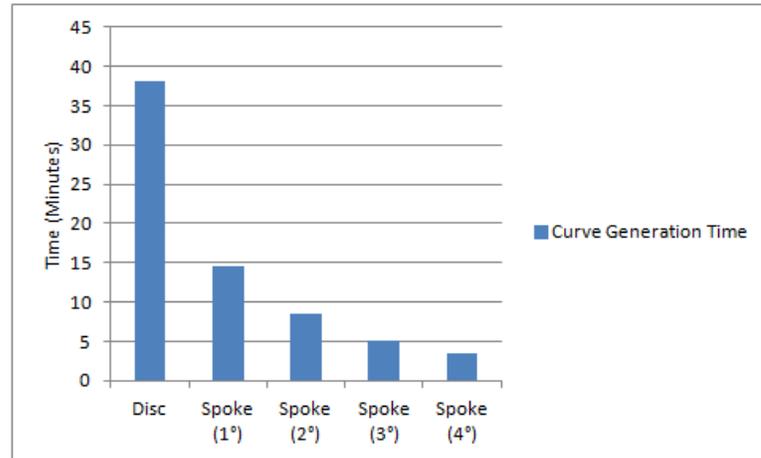


Figure 6.7: Curve generation time for the Epilepsy dataset

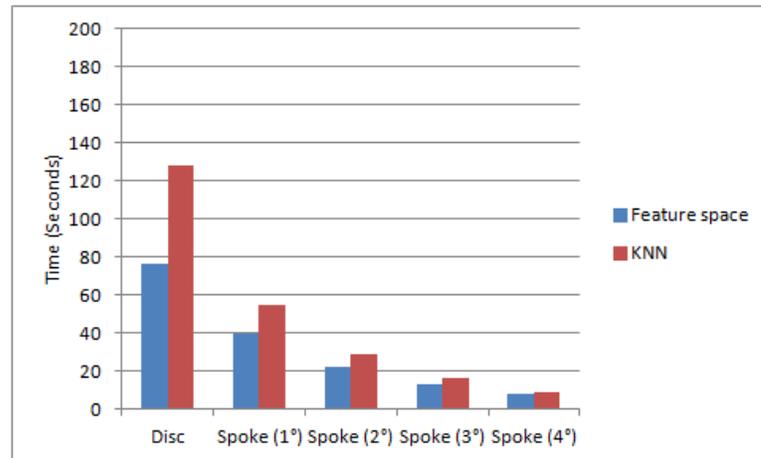


Figure 6.8: Classification time for the Epilepsy dataset

The performance of the proposed point series representation techniques in terms of efficiency is shown in Figures 6.7 to 6.10. Figure 6.7 shows the curve generation time for the Epilepsy dataset, while Figure 6.8 shows the total classification time using TCV for the Epilepsy dataset. Likewise, Figure 6.9 shows the curve generation time for the Musician dataset, while Figure 6.10 shows the total classification time using TCV for the Musician dataset. Note that the time complexity in Figures 6.8 and 6.10 included the Hough signature extraction process (for Feature space classification) and the Dynamic Time Warping similarity measurement process (for KNN classification). From the figures it can be seen that the Spoke-based representation technique ( $4^\circ$ ) was the most efficient with respect to both curve generation and classification for both datasets. This was because a smaller number of points was generated when  $\theta = (4^\circ)$  as

opposed to  $\theta = (3^\circ)$  or less. The Disc-based representation was the least efficient for both datasets because it generates all boundary points.

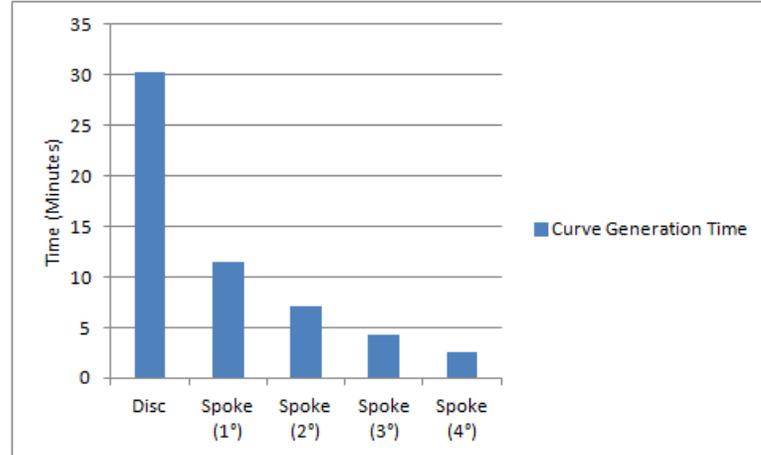


Figure 6.9: Curve generation time for the Musician dataset

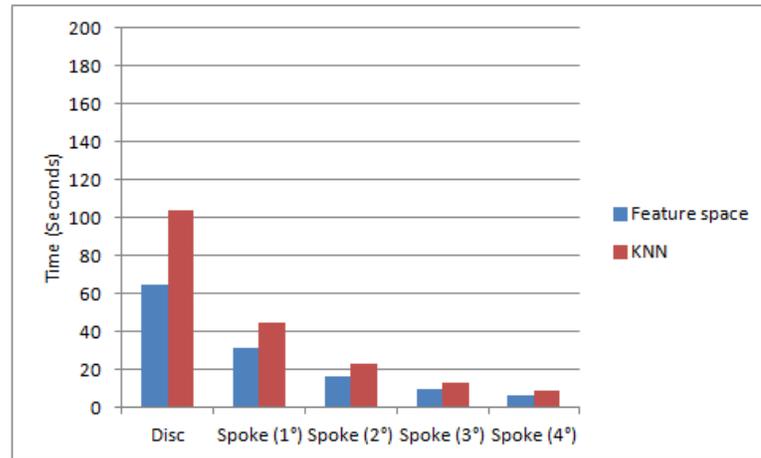


Figure 6.10: Classification time for the Musician dataset

In the case of Spoke-based generation the point series contained fewer points than Disc-based generation. The number of points obtained using the Spoke-based generation was obviously constant for all ventricles (360 point series for  $1^\circ$ , 180 for  $2^\circ$ , 120 for  $3^\circ$  and 90 for  $4^\circ$ ) but varied depending on the size of the ventricles in the case of Disc-based generation (over a thousand points for most ventricles). Consequently, according to Figures 6.7 and 6.9, the Spoke-based representation technique (using  $\theta = 4^\circ$ ) was the most efficient with respect to both curve generation and classification for both the Epilepsy and Musician datasets, while the Disc-based representation was the least efficient for both datasets. However, the performance results using Spoke-based generation with  $\theta = (4^\circ)$  were not good.

For the classification results, those obtained using the Spoke-based representation were, on average, better than those obtained using the Disc-based representation. It is

therefore concluded that the Spoke-based representation produced a better reflection of the shape of the lateral ventricles than in the case of the Disc-based representation.

### 6.3.2 Best Value for $\theta$

With respect to the Spoke-based 3-D image representation technique four different values for  $\theta$  (spoke spacings) were considered:  $\{1^\circ, 2^\circ, 3^\circ, 4^\circ\}$ . Thus four distinct point series were generated in this case. The relationship between classification accuracy and different spoke spacings is presented in Figure 6.11 to 6.14.

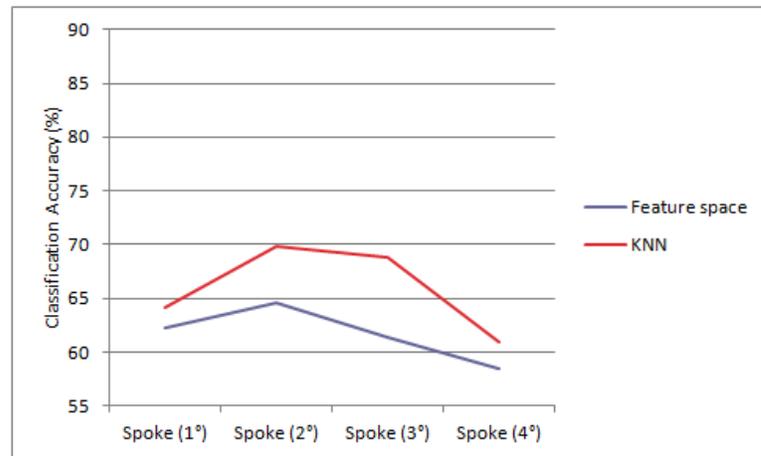


Figure 6.11: The relation between classification accuracy and spoke spacing for the Epilepsy dataset

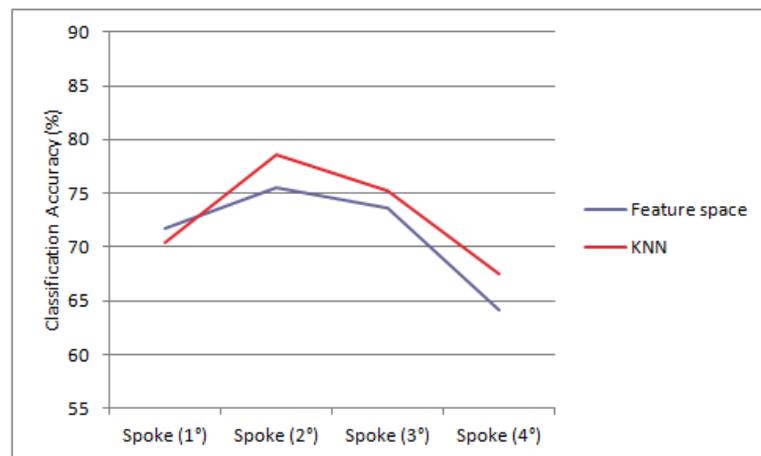


Figure 6.12: The relation between classification accuracy and spoke spacing for the Epilepsy dataset with meta attributes

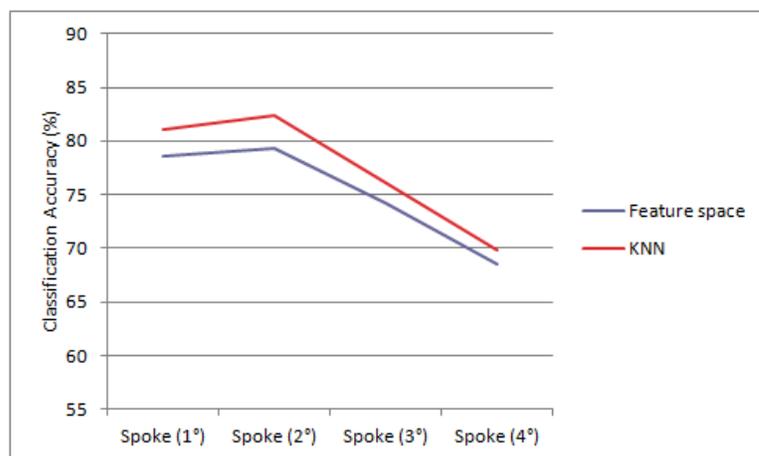


Figure 6.13: The relation between classification accuracy and spoke spacing for the Musician dataset

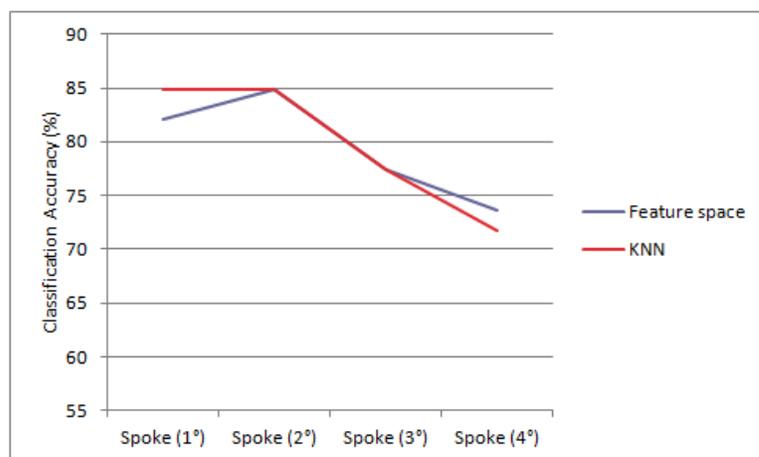


Figure 6.14: The relation between classification accuracy and spoke spacing for the Musician dataset with meta attributes

It can be seen from Tables 6.3 to 6.10 that when using the Spoke-based approach the classification accuracy, regardless of whether meta data was used or not, tended to peak when using a spacing of  $2^\circ$  and then decreased when the spacing increased. It was conjectured that this was because as the representation got coarser details concerning the shape of the ventricles began to be missed, while at lower levels of spacing the amount of detail collected tended to clutter the representation. According to the results, it can be concluded that the spoke spacing of  $2^\circ$  was the best parameter setting for the Spoke-based technique.

### 6.3.3 Feature Space Classification versus KNN Classification

For Feature space classification, the SVM implementation available from the Waikato Environment Knowledge Analysis (WEKA) data mining workbench [54] was adopted.

For KNN classification  $K = 1$  was used. The classification results, in terms of all accuracy, sensitivity and specificity, obtained using KNN classification tended, on average, to be better than those obtained using Feature space classification regardless of whatever a Spoke or Disc-based point series generation was adopted. It was conjectured that this was because the signature extraction process introduced a further level of complexity during which some information concerning the shape of the volume of interest was lost.

### 6.3.4 Use of Meta Attributes

Additional experiments were conducted using two meta attributes: (i) Age and (ii) Gender; the aim here was to determine whether usage of such meta attributes effected the quality of the classification outcomes. Tables 6.7 and 6.8 show the classification results obtained using Feature space classification approach and the age and gender meta attributes for the Epilepsy and Musician datasets respectively, while Tables 6.9 and 6.10 show the results obtained using KNN and the age and gender meta attributes for the Epilepsy and Musician datasets respectively. By comparing the results presented in Tables 6.7 to 6.10 with those presented in Tables 6.3 to 6.6 it can be seen that a significant improvement of effectiveness was recorded when the point series representation was augmented with meta data.

From Tables 6.7 to 6.10 it can be seen that all the results were improved considerably.

Technique	Accu.	Sens.	Spec.
Disc	67.45	71.70	67.07
Spoke ( $x = 1^\circ$ )	72.70	73.58	67.92
Spoke ( $x = 2^\circ$ )	<b>75.47</b>	<b>78.30</b>	<b>69.52</b>
Spoke ( $x = 3^\circ$ )	73.58	75.47	63.81
Spoke ( $x = 4^\circ$ )	64.15	66.04	62.50
Average	70.75	73.02	67.61

Table 6.7: Classification results for the Epilepsy dataset with meta attributes using the Feature space classification

Technique	Accu.	Sens.	Spec.
Disc	78.30	85.85	77.36
Spoke ( $x = 1^\circ$ )	82.08	86.79	81.13
Spoke ( $x = 2^\circ$ )	<b>84.91</b>	<b>90.57</b>	<b>83.02</b>
Spoke ( $x = 3^\circ$ )	77.36	79.25	75.47
Spoke ( $x = 4^\circ$ )	73.58	75.47	73.58
Average	79.25	83.59	78.11

Table 6.8: Classification results for the Musician dataset with meta attributes using the Feature space classification

Technique	Accu.	Sens.	Spec.
Disc	71.04	72.34	66.04
Spoke ( $x = 1^\circ$ )	70.45	74.47	69.81
Spoke ( $x = 2^\circ$ )	<b>78.52</b>	<b>76.67</b>	<b>71.70</b>
Spoke ( $x = 3^\circ$ )	75.18	75.47	67.92
Spoke ( $x = 4^\circ$ )	67.50	70.80	61.32
Average	72.54	73.95	67.36

Table 6.9: Classification results for the Epilepsy dataset with meta attributes using the KNN classification

Technique	Accu.	Sens.	Spec.
Disc	83.96	86.79	83.02
Spoke ( $x = 1^\circ$ )	<b>84.91</b>	<b>90.57</b>	85.85
Spoke ( $x = 2^\circ$ )	<b>84.91</b>	88.68	<b>89.62</b>
Spoke ( $x = 3^\circ$ )	77.36	81.13	75.47
Spoke ( $x = 4^\circ$ )	71.70	77.36	67.92
Average	80.57	84.90	80.38

Table 6.10: Classification results for the Musician dataset with meta attributes using the KNN classification

The efficiency of the proposed point series approach augmented with meta attributes is shown in Figures 6.15 and 6.16. Figure 6.15 shows total classification time using TCV for the Epilepsy dataset and 6.16 shows those for the Musician dataset. Note that the time complexities given in Figures 6.15 and 6.16 included the Hough signature extraction process (for Feature space classification) and Dynamic Time Warping similarity measurement process (for KNN classification). By comparing the results presented in Figures 6.15 and 6.16 with those presented in Figures 6.7 to 6.10 it can be seen that the addition of meta data does not entail a significant computational overhead.

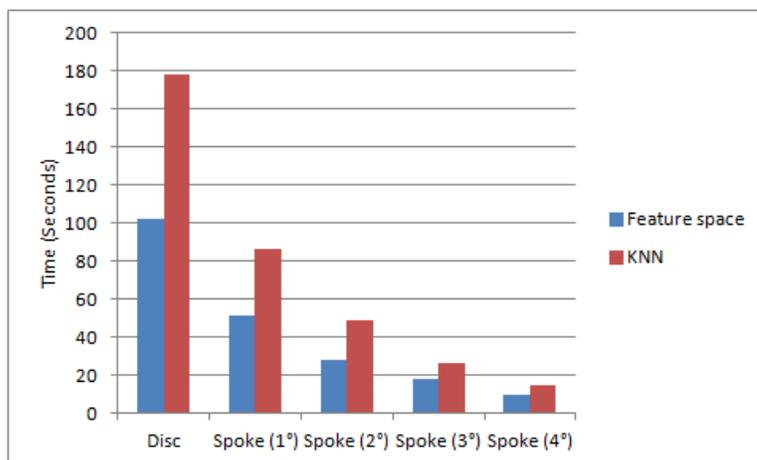


Figure 6.15: Classification time for the Epilepsy dataset with meta attributes

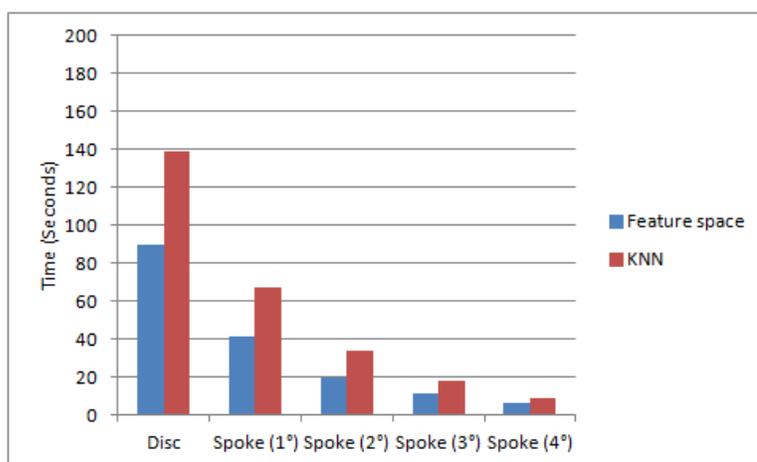


Figure 6.16: Classification time for the Musician dataset with meta attributes

## 6.4 Conclusion

In this chapter an approach to VOIBIC using Disc-based and Spoke-based point series generation coupled with Feature space classification and Direct (KNN) classification has been described. The Hough transform was used to generate signatures describing the curves generated from the proposed point series generation techniques. In the context of direct classification Dynamic Time Warping was used to measure the distance between curves. The reported evaluation indicated that the Spoke-based technique coupled with Direct classification generated the best performance. In the following Chapter an alternative approach to classifying 3-D MRI scans according to a common image feature (VOI), founded on a tree representation, is presented.

## Chapter 7

# Volume of Interest Image Classification Using Oct Tree Based Representation

In this chapter, the Oct-tree based representation for VOIBIC is presented. As already noted in Chapter 2, the application of the techniques to classify 3-D image data to some common object that features across an image set requires the representation of the image objects in question using some appropriate format. The proposed approach presented in this chapter is founded on the concept of Oct-trees and Frequent Subgraph Mining (FSM). In this approach the VOI is represented using an Oct-tree which is then processed further by applying FSM to the tree so as to generate a set frequent sub-graphs that can then be interpreted as features in a feature vector representation to which standard classification techniques can be applied. There are a number of algorithms that can be adopted for FSM, but with respect to the work presented in this thesis the well-known gSpan algorithm [136] was adopted. More specifically a weighted variation of gSpan called gSpan-ATW was used. The input to this algorithm was an Oct-tree, and the output a set of frequently occurring subgraphs together with their occurrence counts. Typically a large number of subgraphs are generated, many of which are redundant (do not serve to discriminate between classes). Feature selection techniques are typically used to reduce the overall number of identified frequent subgraphs. A selected classification mechanisms can then be used to generate a classification model.

Figure 7.1 provides a schematic of the proposed approach. The approach commenced with the extraction of the VOIs (lateral ventricles) explained in Chapter 4. Next, the VOIs were decomposed and represented as Oct-trees. After that, Frequent Subgraph Mining (FSM) was used to discover frequent sub-trees. Then feature vectors were generated and reduced by some selected feature selection techniques. Finally, the classifier was generated and the classification results were obtained.

The remainder of this chapter is organised as follows. Section 7.1 describes the process for generating the desired Oct-tree VOI representation. This is followed in

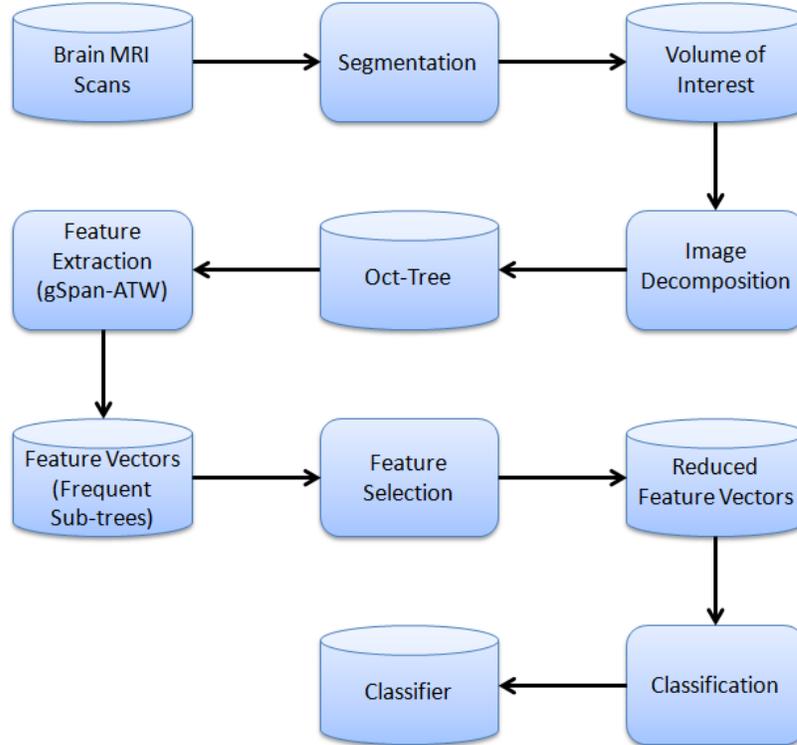


Figure 7.1: Schematic illustrating the graph/tree based approach

Section 7.2 with a description of FSM and the gSpan algorithm in particular. The evaluation of the proposed techniques, including discussion, is then presented in Section 7.3. Finally, the chapter is concluded with a summary in Section 7.4.

## 7.1 Oct-tree Image Representation

The focus of the work presented in this chapter is the representation of VOIs (lateral ventricles) in terms of their boundary using the concept of Oct-tree decomposition. There are many types of image decomposition, common mechanisms use hierarchical data structures; examples for 2-D image data include Quad-trees and Scale Space representations [26]. With respect to 3-D image data, including the work described in this thesis, an Oct-tree representation is typically adopted. Oct-tree is a tree data structure which, in the case of 3-D data, is constructed by repeatedly subdividing a given space into “octants” [99, 68].

In the context of the representation of the VOIs in terms of Oct-tree, the voxels describing each VOI are decomposed into homogeneous sub-regions [38, 41]. The decomposition can be conducted according to a variety of image features such as colour or intensity. With respect to the lateral ventricles a binary encoding was used. The VOI was encapsulated in a minimum bounding box (MBB) The “cubes” in the MBB representation part of the VOI were allocated a “1” (black) and the cubes not included

a“0” (white). The decomposition process entails the recursive decomposition of the VOI into octants. The decomposition continues until either sufficiently homogenous cubes are identified or a user defined “maximum depth” was reached. The result is then stored in the Oct-tree data structure whereby each leaf node represents a cube. Leaf nodes nearer the root of the tree represent larger cubes than nodes further away. Thus the tree is unbalanced in that some leaf nodes will cover larger areas of the VOI than others. It is argued that cubes covering small regions are of greater interests than those covering large regions because they provide more detail. The advantage of this representation is that it maintains information about the relative shape and size of the VOI. An illustration of the Oct-tree generation process is given in Figure 7.2.

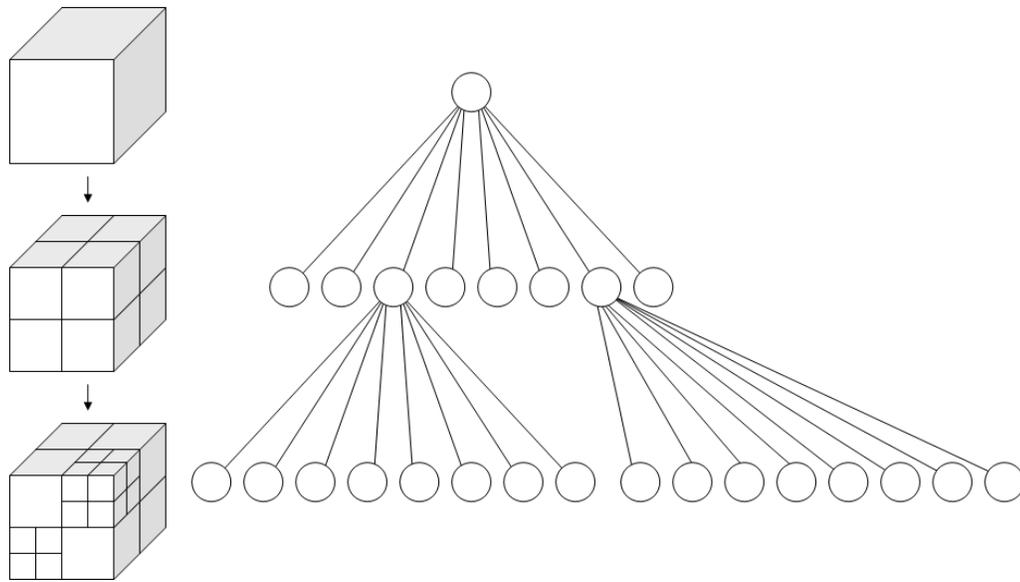


Figure 7.2: Illustration of the Oct-tree hierarchical decomposition process given a VOI surrounded by a MBB

## 7.2 Feature Extraction

Oct-tree are not well suited as input to classifier model generation. The Oct-tree needs to be processed in some way so that a feature vector representation, compatible with classifier model generation, can be formulated. The idea presented here is to apply FSM to the Oct-tree data structure and then used the identified frequent sub-trees to define a feature space. This section presents the process of generating this feature space. Sections 7.2.1 and 7.2.2 provide an overview of graph mining and FSM respectively. Section 7.2.3 then provides a description of gSpan, the adopted FSM algorithm with respect to the work presented in this thesis. Section 7.2.4 then considers the feature vector generation process.

### 7.2.1 Graph Mining

Graph mining is the process of discovering hidden patterns (frequent subgraphs) within graph database. From the literature graph mining can be categorised in terms of transaction graph mining and single graph mining. In transaction graph mining the dataset to be mined comprises a collection of small graphs (transactions). The goal is to discover frequent recurring subgraphs across the dataset. In single graph mining the input of the mining task is one single large graph, and the objective is to find frequent subgraphs which occur within this single graph. Frequent Subgraph Mining (FSM) has demonstrated its advantages with respect to various tasks such as chemical compound analysis [64], document image clustering [9], graph indexing [135], etc.

The straightforward idea behind FSM is to “grow” candidate subgraphs in either a Breadth First Search (BFS) or Depth First Search (DFS) manner (candidate generation), and then determine if the identified candidate subgraphs occur frequently enough in the graph data for them to be considered interesting (support counting). The two main research issues in FSM are thus how to efficiently and effectively: (i) generate the candidate frequent subgraphs and (ii) determine the frequency of occurrence of the generated subgraphs. Effective candidate subgraph generation requires that the generation of duplicate or superfluous candidates is avoided. Occurrence counting requires repeated comparison of candidate subgraphs with subgraphs in the input data, a process known as subgraph isomorphism checking. FSM, in many respects, can be viewed as an extension of Frequent Itemset Mining (FIM) popularised in the context of Association Rule Mining (ARM). Consequently, many of the proposed solutions to addressing the main research issues effecting FSM are based on similar techniques found in the domain of FIM.

It is widely accepted that FSM techniques can be divided into two categories: (i) the Apriori-based approach (also called the BFS strategy based approach) and (ii) the pattern growth approach. These two categories are similar in spirit to counterparts found in ARM, namely the Apriori algorithm [5] and the FPgrowth algorithm [56] respectively. The Apriori-based approach proceeds in a “generate-and-test” manner using a BFS strategy to explore the subgraph lattice of the given database. Therefore, before exploring any  $(k + 1) - subgraphs$ , all the  $k$ -subgraphs should first be explored. For each discovered subgraph  $g$ , this approach extends  $g$  recursively until all the frequent supergraphs of  $g$  are discovered [55]. Pattern growth approaches can use both BFS and DFS strategies, but the latter is preferable to the former because it requires less memory usage. One of the main challenges associate with FIM and FSM is the substantial number of patterns which can be mined from the underlying database. This problem is particularly important in the case of graphs since the size of the output can be extremely large. The significance of FSM with respect to the work described in this thesis is that one of the techniques proposed uses this technique for the purpose

of generating a feature space. The application of FSM algorithms to the datasets described in this work entails a significant computational overhead because of the great number of generated frequent subgraphs. To reduce this overhead a Weighted FSM approach can be applied, the objective being to focus on the identification of those frequent subgraphs that are likely to be the most significant according to some weighting scheme.

## 7.2.2 Frequent Subgraph Mining

The concept of FSM was introduced in Section 7.2.1 above, where it was noted that FSM is an approach to discovering graph structures that occur a significant number of times across a set of graphs. As also noted in Section 7.2.1, there are two separated problem formulations for FSM which can be identified as (i) transaction graph based and (ii) single graph based. Recall that in transaction graph based mining, the input data comprises a collection of relatively small graphs, whereas in single graph based mining the input data comprises a very large single graph. The graph mining based approach with respect to the work in this thesis focuses on transaction graph based mining. The remainder of this section presents a detailed review of FSM. First, the necessary definitions to support the discussion are introduced (the notation used is given in Table 7.1).

**Labelled Graph:** A labelled graph  $G$  is a tuple of the form  $(V, E, l_V, l_E, f_V, f_E)$ , where  $V$  is a set of nodes,  $E \subseteq V \times V$  is a set of edges;  $L_V$  and  $L_E$  are sets of node and edge labels respectively, and  $f_V$  and  $f_E$  are the corresponding functions that define the mappings  $V \rightarrow L_V$  and  $E \rightarrow L_E$ . Note that the edge labels are not included in the definitions used throughout this chapter.

**Subgraph:** Given two graphs  $G_1 (V_1, E_1, L_{V_1}, L_{E_1}, f_{V_1}, f_{E_1})$  and  $G_2 (V_2, E_2, L_{V_2}, L_{E_2}, f_{V_2}, f_{E_2})$ .  $G_1$  is a subgraph of  $G_2$  if  $G_1$  satisfies the following conditions [66].

$$V_1 \subseteq V_2, \forall (v) \in V_1, f_{V_1}(v) = f_{V_2}(v),$$

$$E_1 \subseteq E_2, \forall (u,v) \in E_1, f_{E_1}(u,v) = f_{E_2}(u,v).$$

**Graph Isomorphism:** A graph  $G_1 (V_1, E_1, L_{V_1}, L_{E_1}, f_{V_1}, f_{E_1})$  is isomorphic to an other graph  $G_2 (V_2, E_2, l_{V_2}, l_{E_2}, f_{V_2}, f_{E_2})$ , if and only if a bijection  $\psi : V_1 \rightarrow V_2$  exists such that:

$$\forall (u) \in V_1, f_{V_1}(u) = f_{V_2}(\psi(u)),$$

$$\forall (u), v \in V_1, (u,v) \in E_1 \leftrightarrow (\psi(u), \psi(v)) \in E_2,$$

$$\forall (u,v) \in E_1, f_{E_1}(u,v) = f_{E_2}(\psi(u), \psi(v)).$$

Notation	Description
$D$	A graph database.
$G_i$	A transaction graph such that $G_i \in D$ .
$k - (sub)graph$	A (sub)graph of size $k$ in terms of nodes, edges or paths.
$g_k$	A $k$ -(sub)graph.
$C_k$	A set of subgraph candidates of size $k$ .
$F_k$	A set of frequent $k$ subgraphs.
$ \cdot $	The cardinality of a set.

Table 7.1: Notation used throughout this chapter

The bijection  $\psi$  is an isomorphism between  $G_1$  and  $G_2$ . A graph  $G_1$  is subgraph isomorphic to a graph  $G_2$ , denoted by  $G_1 \subseteq_{sub} G_2$ , if and only if there exists a subgraph  $g$  of  $G_2$  such that  $G_1$  is isomorphic to  $g$  [65].

In the context of transaction graph mining, the FSM aims to discover all the subgraphs whose occurrences in a graph database are over a user defined threshold,  $\sigma$  ( $0 \leq \sigma \leq 1$ ). Formally, given a database  $D$  comprised of a collection of graphs. The occurrence of a subgraph  $g$  in  $D$  is defined by  $\delta_D(g) = \{ G_i \in D \mid g \subseteq_{sub} G_i \}$ . Thus, the support of a graph  $g$  is defined as the fraction of the graphs in  $D$  to which  $g$  is subgraph isomorphic:

$$sup_D(g) = |\delta_D(g)|/|D| \quad (7.1)$$

A subgraph  $g$  is then frequent if and only if  $sup_D(g) \geq \sigma$ . The aim of FSM is thus to find all frequent subgraphs in  $D$ . FSM has been widely studied. A number of FSM algorithms have been proposed including: (i) AGM [66], (ii) FSG [76], (ii) FFSM [65], (iv) SUBDUE [61] or (v) SLEUTH [24]. The weighted FSM algorithm, gSpan-ATW, used with respect to the work presented in this thesis is a gSpan based algorithm [136]. The gSpan algorithm is therefore described in further detail in the following sub-section.

### 7.2.3 gSpan

An outline of the gSpan algorithm is given to Algorithm 3 and 4. The input to Algorithm 3 is a subgraph  $c$  represented by a Depth First Search (DFS) encoding, a minimum support threshold  $\sigma$  and a graph dataset  $D$ . The output is a set of frequent subgraphs  $F$ . Note that the algorithm recursively calls a “subMining” procedure (Algorithm 4) which is used to grow the current frequent subgraphs so as to generate a group of candidate frequent subgraphs. The algorithm works recursively and stops when all the subgraphs, whose appearance count is more than the pre-defined threshold  $\sigma$ , are discovered.

---

**Algorithm 3** The gSpan Algorithm

---

```
1: procedure GSPAN
2:   Sort the labels in  $D$  by their frequency;
3:   Remove infrequent vertices and edges;
4:   Relabel the remaining vertices and edges in descending frequency;
5:    $F_1 \leftarrow \{\text{all frequent 1-edge graphs in } D\}$ ;
6:   Sort  $F_1$  in DFS lexicographic order;
7:    $F \leftarrow \emptyset$ ;
8:   for all  $c \in F_1$  do
9:     subMining( $c, D, \sigma, F$ );
10:     $D \leftarrow D - c$ ;
11:    if  $|D| < \sigma$  then
12:      break;
```

---

---

**Algorithm 4** subMining Procedure

---

```
1: procedure SUBMINING( $c, D, \sigma, F$ )
2:   if  $c \neq \min(c)$  then
3:     return;
4:    $F \leftarrow F \cup \{c\}$ ;
5:    $C_k \leftarrow \emptyset$ ;
6:   Scan  $D$  once, find every edge  $e$  such that  $c$  can be right-most extended to
    $c \cup e, C_k \leftarrow c \cup e$ ;
7:   Sort  $C_k$  in DFS lexicographic order;
8:   for all  $g_k \in C_k$  do
9:     if  $\text{support}(g_k) \geq \sigma$  then
10:      subMining( $g_k, D, \sigma, F$ );
```

---

Normally FSM is naturally computationally expensive because of the candidate generation and support computation processes that are required. The generation of candidate subgraphs needs to be conducted in a non-redundant manner so that the same graph is not generated more than once. Thus graph isomorphism checking is required to remove duplicate graphs. The computation of the support of a graph in the graph database  $D$  also requires subgraph isomorphism checking. The gSpan algorithm uses a canonical labelling and a lexicographical ordering of the input graph in  $D$  to reduce the generation of duplicate graphs. Referring to the subMining procedure shown in Algorithm 4,  $c \neq \min(c)$  (line 2) guarantees that gSpan does not extend any duplicate graph because the candidate generation process adheres to a canonical labelling. Further, the rightmost extension shown at line 6 in the algorithm guarantees that the complete set of frequent subgraphs will be discovered. Although gSpan can achieve a competitive performance compared with other FSM algorithms, its performance degrades considerably when the graph size is relatively large or the graph features many node and/or edge labels. One mechanism for addressing this issue is to use weighted FSM.

To this end a weighted of gSpan was adopted, more specifically gSpan-ATW [136]. The ATW weighing scheme and its consequent incorporation into gSpan are presented in the following two sub-sections.

### The Average Total Weighting (ATW) Scheme

The weighting scheme adopted with respect to the gSpan-ATW algorithm is the Average Total Weighting (ATW) scheme presented in [70], which in turn is inspired by the work of [118]. Given a graph data set  $D = \{G_1, G_2, \dots, G_t\}$ , the weight for a subgraph  $g$  is calculated by dividing the sum of the average weights in the graphs that contain  $g$  with the sum of the average weights across the entire graph data set  $D$ . Thus:

**Definition 7.3.1.** Given a graph dataset  $D = \{G_1, G_2, \dots, G_t\}$  if  $G_i$  is node weighted by assigning  $\{w_1, w_2, \dots, w_k\}$  to a set of nodes  $\{v_1, v_2, \dots, v_k\}$  respectively, then the average weight associated with  $G_i$  is defined as:

$$W_{avg}(G_i) = \frac{\sum_{j=1}^k w_j}{k_i} \quad (7.2)$$

Where  $w_j$  can be determined using an appropriate weighting function described as follows:

$$w_j = \frac{occ(v_j)}{\sum_{1 < i < t} size(G_i)} \quad (7.3)$$

Where  $occ(v_j)$  denotes the number of times  $v_j$  occurs in  $D$ , and  $size(G_i)$  denotes the size of  $G_i$  in terms of the number of nodes in  $G_i$ . Thus, the total weight of  $D$  is further defined as:

$$W_{sum}(D) = \sum_{i=1}^t W_{avg}(G_i) \quad (7.4)$$

Using both equation (7.2) and (7.4), the weight of a subgraph can be calculated using equation (7.5).

**Definition 7.3.2.** Given a graph data set  $D = \{G_1, G_2, \dots, G_t\}$  and an arbitrary graph  $g$ , let the set of graphs where  $g$  occurs equal  $\delta_D(g)$ . Then, the weight of  $g$  with respect to  $D$  is:

$$W_D(g) = \frac{\sum_{G_i \in \delta_D(g)} W_{avg}(G_i)}{W_{sum}(D)} \quad (7.5)$$

$W_D(g)$  is used to quantify the actual importance of each discovered subgraph  $g$  in a graph database. According to (7.1), the weighted support of a graph  $g$  is then defined as the product of the support of  $g$ :

$$wsup_D = W_D(g) \cdot sup_D(g) = \frac{W_D(g) \cdot |\delta_D(g)|}{|D|} \quad (7.6)$$

**Definition 7.3.3.** A subgraph  $g$  is weighted frequent with respect to  $D$ , if  $wsup_D(g) \geq \tau$ , where  $0 < \tau \leq 1$  is a weighted support threshold.

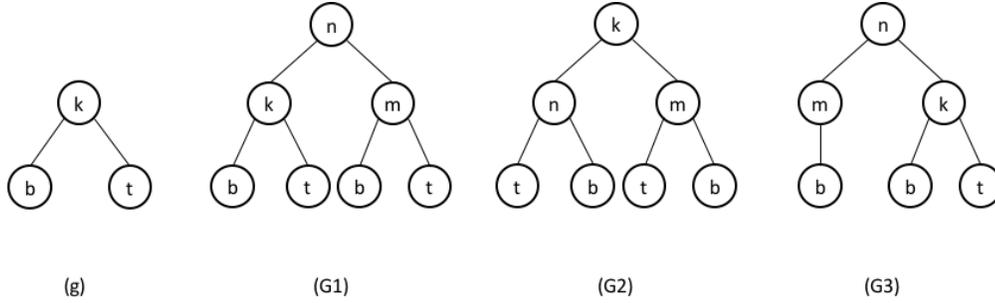


Figure 7.3: Example dataset for illustrating the process of calculating weights using the ATW scheme

*Example:* Considering the graph data set  $D = \{G_1, G_2, G_3\}$  shown in Figure 7.3, where the letter inside each node indicates the node label. Given a node with a label ‘b’ in the candidate subgraph  $g$  (in Figure 7.3),  $occ(b) = 6$ , and  $\sum_{i=1}^3 size(G_i) = 20$ . Thus  $w_D(b) = 6/20 = 0.3$ . Similarly, for nodes with labels ‘t’, ‘n’, ‘m’ and ‘k’ in  $D$ , the weights of the nodes are  $w_D(t) = 5/20 = 0.25$ ,  $w_D(n) = 3/20 = 0.15$ ,  $w_D(m) = 3/20 = 0.15$  and  $w_D(k) = 3/20 = 0.15$  respectively. Given a subgraph  $g$ , which occurs in  $G_1$  and  $G_3$ ,  $W_{avg}(G_1) = \frac{0.15+0.15+0.15+0.25+0.3+0.3+0.25}{7} \approx 0.2214$ ,  $W_{avg}(G_2) = \frac{0.15+0.15+0.15+0.25+0.3+0.25+0.3}{7} \approx 0.2214$ ,  $W_{avg}(G_3) = \frac{0.15+0.15+0.15+0.3+0.3+0.25}{6} \approx 0.2167$ .

Thus:

$$W_{sum}(D) = 0.2214 + 0.2214 + 0.2167 \approx 0.6595, W_D(g) = \frac{0.2214+0.2167}{0.6595} \approx 0.6643, wsup_D = 2/3 \times W_D(g) \approx 0.4429.$$

### The gSpan-ATW Algorithm

The ATW weighting scheme was incorporated into the gSpan algorithm to produce gSpan-ATW [136]. To do this the procedure ‘subMining’, Algorithm 4 described in Section 7.2.3 needed to be replaced with an alternative procedure ‘subMining-ATW’.

This alternative procedure is presented in Algorithm 5. In Algorithm 5, a weighted support threshold  $\tau$  is introduced to replace the threshold  $\sigma$  used in Algorithm 3.

The subMining-ATW algorithm operates in a similar manner to the subMining algorithm discussed above; level by level, following a “generate, calculate support and prune” loop. A candidate subgraph,  $g$ , is considered to be frequent if its weighted support,  $wsup_D(g)$ , is greater than some user specified threshold,  $\tau$ ; otherwise it is pruned. Note that, as in the case of the  $\sigma$  threshold, the lower the value of  $\tau$  the greater the number of frequent subgraphs that will be identified.

---

**Algorithm 5** subMining-ATW Procedure

---

```

1: procedure SUBMINING-ATW( $(c, D, \sigma, F)$ )
2:   if  $c \neq \min(c)$  then
3:     return;
4:   if  $W_d(c) \times sup_D(c) \geq \tau$  then
5:      $F \leftarrow F \cup \{c\}$ ;
6:   else
7:     return;
8:    $C_k \leftarrow \emptyset$ ;
9:   Scan  $D$  once, find every edge  $e$  such that  $c$  can be right-most extended to
    $c \cup e, C_k \leftarrow c \cup e$ ;
10:  Sort  $C_k$  in DFS lexicographic order;
11:  for all  $g_k \in C$  do
12:    if  $W_D(g_k) \times sup_D(g_k) \geq \tau$  then
13:      subMining-ATW( $g_k, D, \tau, F$ );

```

---

### 7.2.4 Feature Vector Generation

Using the proposed gSpan-ATW algorithm, the identified frequent subgraphs (or sub-trees) describe, in terms of size and shape, some part of a VOI that occurs regularly across the dataset. These are then used to form the fundamental elements of a feature space. In this context a feature space is an  $N$  dimensional space where  $N$  is equivalent to the number of features identified. Using this feature space each image (VOI) can be described in terms of a feature vector of length  $N$ , with each element having a value equal to the frequency of that feature. The feature vectors for each MRI brain scan with respect to the work presented in this chapter comprised the occurrence count for each identified frequent subgraph. As in the case of the work presented in the previous chapter this could be augmented with subject age and gender information.

## 7.3 Evaluation

This section describes the evaluation of the proposed graph based approach to VOIBIC. A number of experiments were conducted to measure the effectiveness and efficiency of the proposed approach. The objectives of the evaluation were as follows:

1. To compare the usage of two alternative classification model generators, Support Vector Machine (SVM) and Decision Tree, in the context of the proposed Oct-tree representation and classification effectiveness.
2. To identify the most appropriate value for the minimum support threshold  $\tau$  for use with respect to the gSpan-ATW algorithm.
3. Whether the use of additional meta data, age and gender, would effect the quality of the classification.
4. To compare the run time efficiency of the proposed approach.

The image sets used for the evaluation were again the Epilepsy and Musician datasets introduced in Chapter 3.

With respect to the gSpan-AWT algorithm four different minimum support thresholds for  $\tau$  were used to define frequent subgraphs ( $\{20\%, 30\%, 40\%, 50\%\}$ ). As a result, four sets of feature vectors were generated and used as inputs for each classifier. To reduce the number of features (frequent sub-trees) Chi-square feature selection [81], available within the Waikato Environment Knowledge Analysis (WEKA) data mining workbench [54], was used. For the evaluation Ten-Fold Cross Validation (TCV) was used, this results presented in the following subsections are average results across all ten folds. All the experiments were conducted using a 2.9 GHz Intel Core i7 with 8GB RAM on OS X (10.9) operating system. The evaluation results obtained with respect to each of the above objectives are discussed in Section 7.3.1 to 7.3.3 below.

$\tau$ (%)	SVM			Decision Tree		
	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.
20	68.53	70.40	68.53	67.83	70.58	68.67
30	<b>72.34</b>	<b>75.67</b>	<b>70.45</b>	<b>70.45</b>	<b>74.28</b>	<b>73.20</b>
40	70.45	72.34	69.23	65.87	70.47	65.40
50	62.28	66.96	60.67	62.80	67.05	64.15
Average	68.40	71.34	67.22	66.74	70.60	67.86

Table 7.2: Classification results for the Epilepsy dataset using SVM and Decision Tree

### 7.3.1 Support Vector Machine classification versus Decision Tree Classification

As noted above for the comparison of the effectiveness of SVM classification versus Decision Tree classification experiments were conducted using the implementations provided within the Waikato Environment Knowledge Analysis (WEKA) data mining workbench [54]. Tables 7.2 and 7.3 show the classification results obtained using the proposed Oct-tree based approach to VOIBIC for the Epilepsy and Musician datasets

$\tau$ (%)	SVM			Decision Tree		
	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.
20	85.77	<b>90.02</b>	<b>81.53</b>	82.08	<b>86.79</b>	77.36
30	<b>86.32</b>	87.74	80.19	<b>83.02</b>	85.85	<b>80.19</b>
40	81.60	84.91	78.30	78.30	81.13	75.47
50	75.00	80.19	69.81	72.17	77.36	66.98
Average	82.17	85.72	77.46	78.89	82.78	75.00

Table 7.3: Classification results for the Musician dataset using SVM and Decision Tree

respectively (best results highlighted in bold font). Note that the abbreviations ‘Accu.’, ‘Sens.’ and ‘Spec.’ used in the Tables in this Chapter indicate classification accuracy, sensitivity and specificity respectively.

From Table 7.2 (Epilepsy) it can be seen that, with respect to Oct-tree classification, the best classification accuracy, sensitivity and specificity were obtained when using  $\tau = 30\%$  for both SVM and Decision Tree. From Table 7.3 (Musician), the best classification accuracy for both SVM and Decision Tree were also obtained when using  $\tau = 30\%$  while the best classification sensitivity for both SVM and Decision Tree were obtained when using  $\tau = 20\%$ . For the musicians dataset the best classification specificity for SVM was obtained when using  $\tau = 20\%$  while those for Decision Tree was obtained when using  $\tau = 30\%$ . From the average results presented in the Tables 7.2, in terms of effectiveness, for the Epilepsy data set the two classification models both operated in a similar manner; there was little to choose between them. For the musicians data set SVM outperformed the decision tree classification model.

### 7.3.2 Best Value of Minimum Support Threshold

Referring back to Tables 7.2 and 7.3, with respect to both classification methods, it can be seen that when using  $\tau = 20\%$  and  $\tau = 30\%$ , regardless of the dataset used, the results were not significantly different in terms of classification accuracy, sensitivity and specificity. However, the results then tended to decrease when the  $\tau$  was increased beyond 30%. This was because more frequent subgraphs were identified with low support thresholds. It was likely that as the support threshold increases, significant subgraphs were not discovered by the gSpan-AWT algorithm. The results obtained therefore indicate that the best minimum support thresholds values to be used for  $\tau$ , with respect to the proposed oct-tree based VOIBC approach, was between 20% and 30%.

### 7.3.3 Use of Meta Attributes

Additional experiments were conducted using the two meta attributes used previously: (i) age and (ii) gender; the aim here was to determine whether usage of such meta

attributes effected the quality of the classification outcomes. Tables 7.4 and 7.5 show the classification results obtained using the Oct-tree classification approach for the Epilepsy and Musician datasets augmented with meta attributes (the best results are shown in bold font).

$\tau$ (%)	SVM			Decision Tree		
	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.
20	77.36	<b>82.08</b>	72.64	75.57	80.19	70.75
30	<b>78.30</b>	81.13	<b>75.47</b>	<b>76.42</b>	<b>81.13</b>	<b>71.70</b>
40	75.47	80.19	70.75	70.28	76.42	64.15
50	67.45	72.64	62.26	64.15	68.87	59.43
Average	74.65	79.01	70.28	71.61	76.65	66.51

Table 7.4: Classification results for the Epilepsy dataset augmentation and using SVM and Decision Tree

$\tau$ (%)	SVM			Decision Tree		
	Accu.	Sens.	Spec.	Accu.	Sens.	Spec.
20	<b>86.02</b>	88.85	<b>83.19</b>	<b>85.37</b>	<b>92.50</b>	<b>78.57</b>
30	85.37	<b>92.50</b>	78.57	84.15	90.00	<b>78.57</b>
40	82.93	87.50	78.57	81.71	87.50	76.19
50	76.83	77.50	76.19	80.00	77.36	69.05
Average	82.79	86.59	79.13	82.81	86.84	75.60

Table 7.5: Classification results for the Musician dataset augmentation and using SVM and Decision Tree

From Table 7.4 it can be seen that the best classification accuracy, sensitivity and specificity results were obtained using Decision Tree classification and  $\tau = 30\%$ . The best classification accuracy and specificity for SVM were also obtained when using  $\tau = 30\%$  while the best classification sensitivity was obtained using  $\tau = 20\%$ . From Table 7.5, all the best classification accuracy, sensitivity and specificity for Decision Tree were obtained when  $\tau = 20\%$ . Likewise, the best classification accuracy and specificity for SVM were obtained when  $\tau = 20\%$ . However the best classification sensitivity for SVM was obtain when  $\tau = 30\%$ .

Comparing the results presented in Tables 7.4 and 7.5, using meta data, with those presented previously in Tables Tables 7.2 and 7.3, where meta data was not used, it can be seen that the usage of meta data improved the performance of the classifiers. This was especially so in the case of the Epilepsy dataset. The results presented in Tables 7.4 and 7.5 also corroborate the conclusions draw from the earlier results: (i) the most appropriate value for  $\tau$  is between 20%and 30% and (ii) SVM tended to out perform decision trees in terms of classification effectiveness.

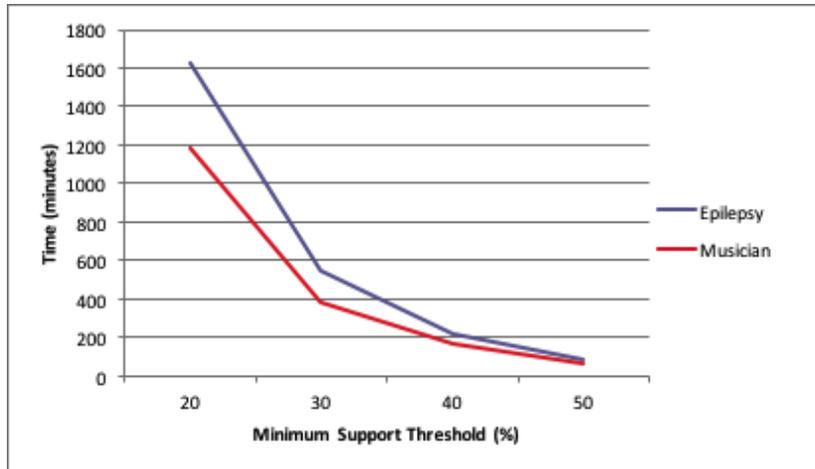


Figure 7.4: Run time complexity of Oct-tree generation and frequent subgraph mining process using gSpan

### 7.3.4 Run Time Efficiency

In terms of the run time efficiency of the VOIBIC approach based on the proposed Oct-tree representation techniques a sequence of plots are given in Figures 7.4 to 7.8. Figure 7.4 shows Oct-tree generation time (minutes), for both the Epilepsy and Musician datasets, against a range  $\tau$  threshold values. Note that the Oct-tree generation time includes FSM (using gSpan) process. Figures 7.5 and 7.6 show the total TCV classification time for the Epilepsy and Epilepsy+ (with meta data) datasets using SVM and Decision Tree classification respectively. Likewise, Figures 7.7 and 7.8 show the total TCV classification time for the Musician and Musicians+ (with meta data) dataset using SVM and Decision Tree classification respectively.

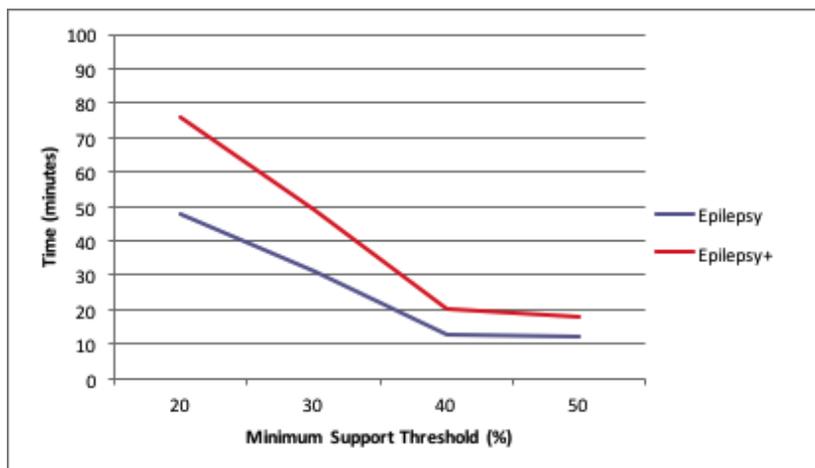


Figure 7.5: Run time complexity of image classification for the Epilepsy dataset using SVM classification

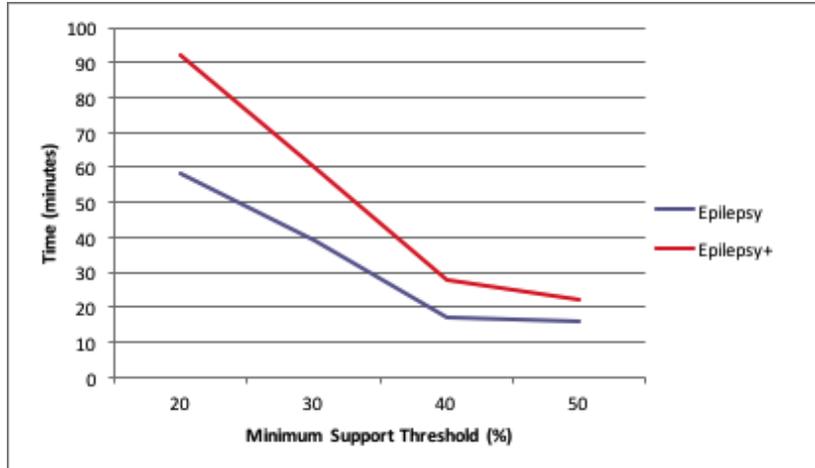


Figure 7.6: Run time complexity of image classification for the Epilepsy dataset using Decision Tree classification

From Figure 7.4 it can be seen that the run time for Oct-tree generation and FSM process of Epilepsy dataset was slower than that of Musician dataset. The reason why the runtime in the case of the Epilepsy dataset was longer than that for Musician dataset was simply because the Epilepsy dataset was larger (210 volumes versus 160 volumes). On the other hand, the run time for Oct-tree generation and FSM process was least efficient when using  $\tau$  (minimum support threshold) = 50% and best efficient when using  $\tau = 20\%$ . This was because the number of subgraphs discovered from the process decreased when  $\tau$  increased.

From Figures 7.5 to 7.8 it can be seen that, as expected, total classification time, with respect to the proposed Oct-tree representation techniques, was the most efficient when using  $\tau = 50\%$  for both classification methods and both dataset; while when using  $\tau = 20\%$  was the least efficient. This was because the number of identified frequent subgraphs decreased as  $\tau$  increased. Of course, as noted above, lower values of  $\tau$  are more effective. The reason why the runtime in the case of the Epilepsy dataset was longer than that for Musician dataset was simply because the Epilepsy dataset was larger (210 volumes versus 160 volumes). Inspection of the figures also indicates that the SVM classification mechanism performed slightly faster than decision tree mechanism for both the Epilepsy and Musician datasets. Therefore, according to the reported results, in this sub-section and the earlier sub-sections, it can be concluded that SVM was a better classification mechanism than DT in terms of both effectiveness and efficiency.

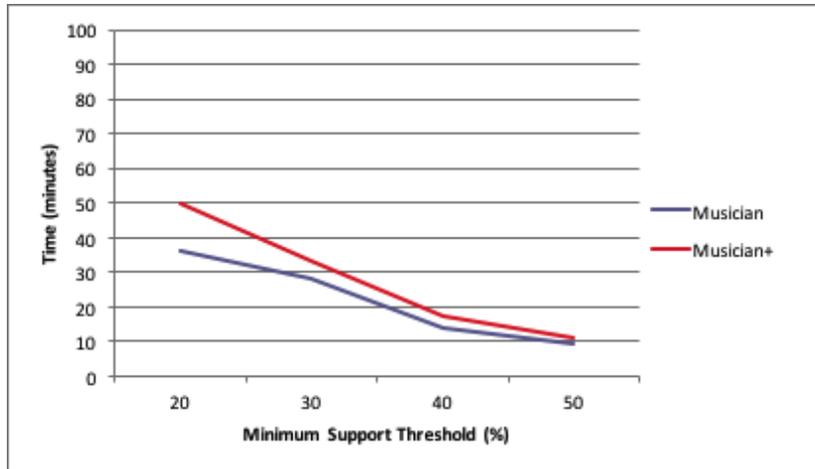


Figure 7.7: Run time complexity of image classification for the Musician dataset using SVM classification

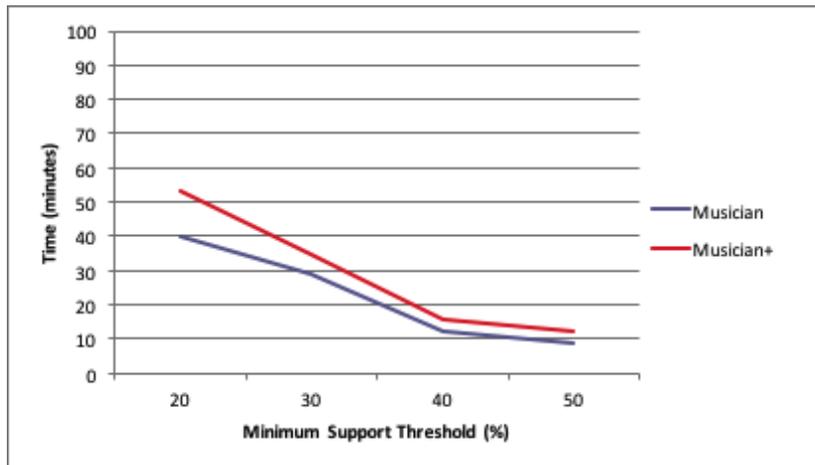


Figure 7.8: Run time complexity of image classification for the Musician dataset using Decision Tree classification

## 7.4 Conclusion

In this chapter an approach to VOIBIC, using an Oct-tree representation, and the use of FSM has been proposed. The approach used a hierarchical decomposition whereby each MRI scan was decomposed into hierarchy of “cubes” which could then be represented using an Oct-tree structure. A FSM mechanism was then applied so that subgraphs that occurred frequently across the entire image collection could be identified. The frequent subgraphs were considered to describe a feature space; as such the input image could be translated, according to this feature space, into a set of feature vectors (one per image). A feature selection mechanism was applied to the feature vectors so as to select the most significant features to which standard classification techniques could be applied. The reported evaluation indicated that high classification accuracy results

were obtained when using gSpan-ATW minimum support threshold ( $\tau$ ) values of 20% and 30%. In the following chapter, this thesis is discussed with comparison among of the proposed techniques.

# Chapter 8

## Discussion

This penultimate chapter provides a discussion of the work presented in this thesis. The main discussion points are directed at: (i) a comparison of the proposed approaches in terms of earlier evaluation results, (ii) a statistical significance comparison and (iii) a comparison with the previous work. The chapter is organised as follows. Section 8.1 presents the comparison among the proposed approaches in terms of the evaluation results presented earlier in this thesis; results obtained from experiments undertaken in terms of both classification effectiveness and efficiency. The following section, Section 8.2, then presents the overall results obtained from statistical significance testing. Finally, the comparison with the previous related work is then given in Section 8.3.

### 8.1 Comparison in Terms of Earlier Evaluation Results

The comparison of the proposed VOIBIC representations with respect to the evaluation results presented earlier in this thesis was undertaken in terms of: (i) classification effectiveness and (ii) efficiency. In terms of effectiveness the results reported earlier were generated using a variety of parameter settings for each approach, and in some cases variations of the approach under consideration. For the overall comparison reported in this section the best performing parameter settings and/or variations with respect to each approach was used (so as to consider each technique to its best advantage). The comparison is also divided into two parts: (i) without augmentation and (ii) with augmentation.

Tables 8.1 and 8.2 summarise the “best” classification results of obtained for the Epilepsy and musicians datasets respectively. In the tables the acronyms ‘SMB’, ‘DB’, ‘SB’ and ‘TB’ refer to the Statistical Metric Based, Disc Based, Spoke Based, and Tree Based approaches respectively (recall that the point series based technique was split into two techniques, ‘DB’ and ‘SB’). Thus four different approaches are presented. The ‘Accu.’, ‘Sens.’ and ‘Spec.’ abbreviations indicate classification accuracy, sensitivity and specificity respectively as before. Best results are indicated in bold font.

From Table 8.1 it can be seen that, with respect to the Epilepsy dataset, the overall

Technique	Accu.	Sens.	Spec.
SMB	60.63	63.56	57.16
DB	62.20	67.50	57.14
SB	69.81	71.70	67.92
TB	<b>72.34</b>	<b>75.67</b>	<b>70.45</b>

Table 8.1: Best classification effectiveness results for Epilepsy dataset (without augmentation)

classification accuracies obtained using the four proposed approaches were over 60%. All the classification accuracy, sensitivity and specificity values obtained using the Tree based approach were higher than the other three approaches considered. The lowest classification accuracy and sensitivity values were obtained using the Statistical metrics based approach, and the lowest classification specificity was obtained using Disc-based approach.

Technique	Accu.	Sens.	Spec.
SMB	69.89	71.77	68.00
DB	77.36	81.13	75.47
SB	82.39	<b>88.68</b>	79.25
TB	<b>86.32</b>	87.74	<b>80.19</b>

Table 8.2: Best classification effectiveness results for Musician dataset (without augmentation)

From Table 8.2 it can be seen that, with respect to the Musicians dataset, the overall classification accuracies obtained using the four proposed approaches were over 67%. The classification accuracy and specificity obtained using the Tree based approach were the best compared to the other three approaches, while the classification sensitivity obtained using the Spoke-based approach was the best compared to the other three approaches. The lowest classification accuracy, sensitivity and specificity values were obtained using the Statistical metrics based approach.

Technique	Accu.	Sens.	Spec.
SMB	68.59	70.50	68.17
DB	71.04	72.34	66.04
SB	<b>78.52</b>	76.67	71.70
TB	78.30	<b>81.13</b>	<b>75.47</b>

Table 8.3: Best classification effectiveness results for Epilepsy dataset (with augmentation)

Tables 8.3 and 8.4 present a summary of the best classification effectiveness results using augmented data with respect to the epilepsy and musicians data sets. From

Table 8.3 the results indicate that, with respect to the augmented Epilepsy dataset, the overall classification accuracies obtained using the four proposed approaches were over 68%. The classification sensitivity and specificity values obtained using the Tree based approach were better than the other three approaches, while the classification accuracy obtained using the Spoke-based approach was the best compared to the other three approaches. The lowest classification accuracy and sensitivity were obtained using the Statistical metric based approach, and the lowest classification specificity was obtained using Disc-based approach.

Technique	Accu.	Sens.	Spec.
SMB	70.36	77.00	64.05
DB	83.96	86.79	83.02
SB	84.91	<b>90.57</b>	<b>85.85</b>
TB	<b>86.02</b>	88.85	83.19

Table 8.4: Best classification effectiveness results for Musician dataset (with augmentation)

From Table 8.4 it can be seen that, with respect to the augmented Musicians dataset, the results indicated that the overall classification accuracies obtained using the four proposed approaches were over 70%. The classification accuracy obtained using the Tree based approach was the best compared to the other three approaches while the classification sensitivity and specificity obtained using the Spoke-based approach were the best compared to the other three approaches. All the lowest classification accuracy, sensitivity and specificity results were obtained using the Statistical metrics based approach.

For the overall comparison in terms of efficiency (run time) the best performing parameters and/or variations with respect to each technique were again used; and, as before, the comparison was conducted with and without augmented data. Recall that the recorded runtimes (seconds) were obtained over all ten runs of TCV where each run includes: (i) representation generation, (ii) classification model generation, and (iii) classification model testing. All the experiments were conducted using a 2.9 GHz Intel Core i7 with 8GB RAM on OS X (10.9) operating system.

The efficiency result summaries are presented in Figures 8.1 to 8.2. In the figures, as before, ‘SMB’, ‘DB’, ‘SB’ and ‘TB’ indicate the Statistical Metric Based, Disc Based, Spoke Based and Tree Based approaches respectively. In the figures the augmented and non-augmented versions of the Epilepsy and Musicians data sets are indicated by ‘Epilepsy’ and ‘Epilepsy+’, and ‘Musician’ and ‘Musician+’ respectively. From the figures, it can be seen that Statistical metrics based approach was the most efficient for all datasets (The run time was less than one minute) while the Tree based approach was the least efficient for all datasets.

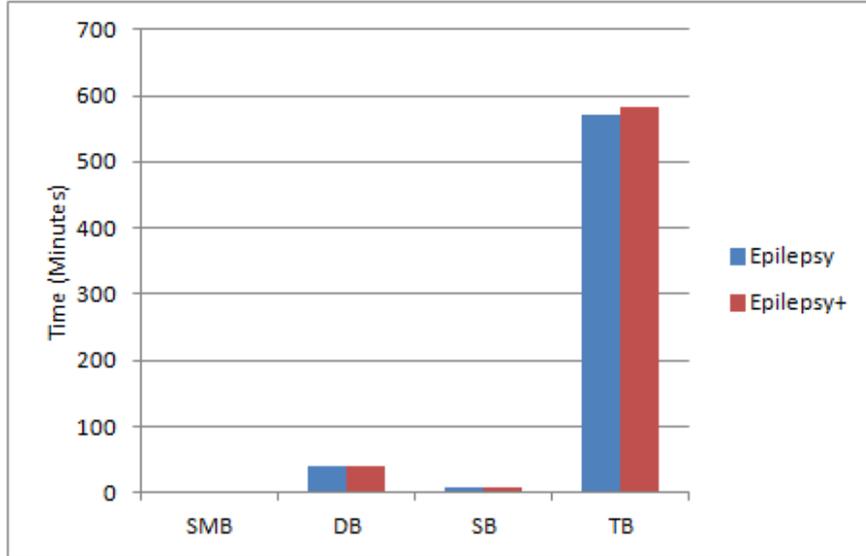


Figure 8.1: Run time complexity for the classification process for the Epilepsy dataset (with and without augmentation)

Thus, with respect to the work presented in this thesis, promising classification results were produced. From the foregoing it can be seen that there is a trade-off between effectiveness and efficiency. If we wished to maximise both effectiveness and efficiency it would be possible to argue that Spoke-based representation approach produced the best results. However, in a medical context, such as that considered as a focus in this thesis, effectiveness outranks efficiency. In the next section a statistical significance comparison of the proposed approaches is presented.

## 8.2 Statistical Significance Comparison

This section presents the results obtained from conducting statistical significance testing. Given any competition, at some level of granularity, one of the competitors will always come first, the question is whether this outcome is statistically significant or not? If ten people run down a corridor and get to the other end within seconds of each other it might not be statistically significant to say that the winner was the fastest it might simply be chance; if, on the other hand, the winner gets to the end of the corridor thirty seconds ahead of the rest then it might be statistically significant to say that he/she was the fastest.

There are a variety of mechanisms that can be used to establish statistical significance. The mechanism used here is the Friedman test [32, 45]. The Friedman test statistic,  $\chi_F^2$  is based on the average ranked (AR) performances of the classification techniques applied to each dataset, and is calculated as follows:

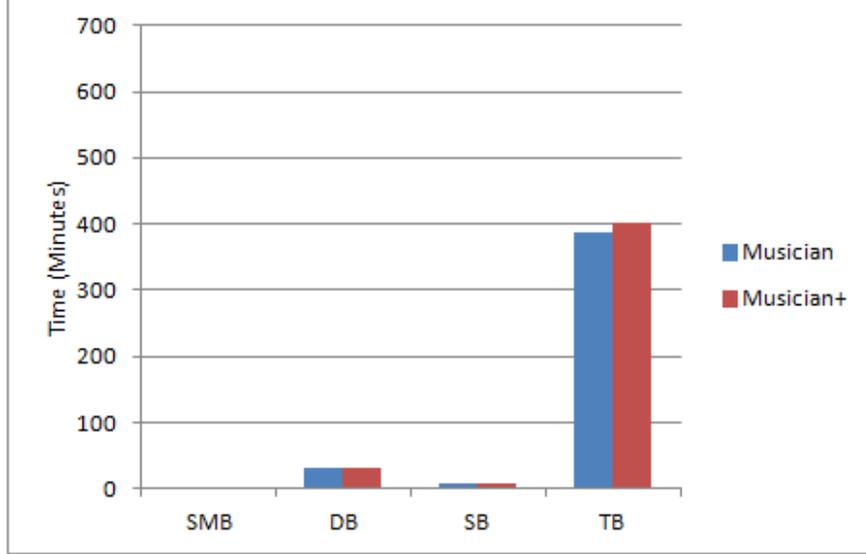


Figure 8.2: Run time complexity for the classification process for the Musician dataset (with and without augmentation)

$$\chi_F^2 = \frac{12N}{K(K+1)} \left[ \sum_{j=1}^K AR_j^2 - \frac{K(K+1)^2}{4} \right] \quad (8.1)$$

where: (i)  $AR_j = \frac{1}{N} \sum_{i=1}^N r_i^j$ , (ii)  $N$  denotes the number of datasets used in the study, (iii)  $K$  is the total number of classifiers considered and (iv)  $r_i^j$  is the rank of classifier  $j$  with respect to dataset  $i$ .  $\chi_F^2$  is distributed according to the Chi-square distribution with  $K - 1$  degrees of freedom. If the value of  $\chi_F^2$  is large enough, then the null hypothesis that there is no difference between techniques can be rejected. The Friedman statistic is well suited for comparing the significance of the results obtained with respect to the proposed VOIBIC representations.

If the null hypothesis can be rejected a post hoc Nemenyi test [32] can be applied to highlight significant differences between the individual representations. The Nemenyi post hoc test states that the performances of two or more classifiers are significantly different if their average ranks differ by at least a Critical Difference (CD), given by:

$$CD = q_{\alpha, \infty, K} \sqrt{\frac{K(K+1)}{12N}} \quad (8.2)$$

Note that in Equation 8.2, the value  $q_{\alpha, \infty, K}$  is based on the Studentised range statistic [32].

For the purpose of the evaluation AUC values were used. Table 8.5 reports the AUCs for all four techniques when applied to the two datasets augmented with age and gender. In the table, the techniques achieving the highest AUC with respect to each dataset, and the overall highest ranked technique, are indicated in bold font. The numbers in the parentheses indicate the average rank of each technique. The Friedman

test statistic and corresponding p-value are also shown. From the table it can be seen that the Tree Based approach (TB) has the highest Friedman score (average rank - AR). The AR of the Point series approach is statistically worse than the AR for the Tree based approach at the 5% critical difference level ( $\alpha = 0.05$ ). As the  $\chi_F^2$  values were all significant ( $p < 0.005$ ) the null hypothesis that there is no difference between the techniques could be rejected and a post hoc Nemenyi test applied to each class distribution.

Figure 8.3 shows the critical difference diagram for the data presented in Table 8.5. Note that this is a modified version of the Demsar 2006 significant diagram [78]. This figure shows the classification approaches listed in ascending order of ranked performance on the y-axis, and the image classification approaches' average rank across two datasets displayed on the x-axis. The diagram displays the ranked performances of the classification techniques, along with the critical difference tail, to highlight any techniques which are significantly different to the best performing techniques. The CD value for the figure was calculated as per Equation 8.2. The critical difference diagram clearly shows that the TB approach is the best performing classification approach with an AR value of 1.0; however, this result is not significant with respect to SB and DB, while it is significant with respect to SMB.

Friedman test statistic = 6.00 ( $p < 0.005$ )			
Approach	Epilepsy (AUC.)	Musician (AUC.)	Average Rank (AR)
SMB	65.62 (4)	66.68 (4)	4
DB	68.12 (3)	78.59 (3)	3
SB	76.95 (2)	85.76 (2)	2
TB	<b>77.30 (1)</b>	<b>86.10 (1)</b>	<b>1</b>

Table 8.5: Area under the receiver operating characteristic Curve (AUC) results

### 8.3 Comparison with Previous Work

As noted in Section 3.2 of Chapter 3 there has been some previous work directed at the epilepsy and musician datasets. Notably, that of Elsayed et al. [40, 37]. Although it should be noted that in [37] the data was considered in terms of 2-D and focused on the corpus callosum (an alternative feature found in MRI brain scans to the lateral ventricles). It is possible to make some comparisons. The results are shown in Table 8.6 where the best AUC results with respect to the work presented in this thesis are listed alongside the best AUC results reported in [37].

From the table it can be seen that in both cases better performance are recorded with respect to the musician dataset than the epilepsy dataset. This is because epilepsy is clearly a more complex conditions than musician ability. The results reported in

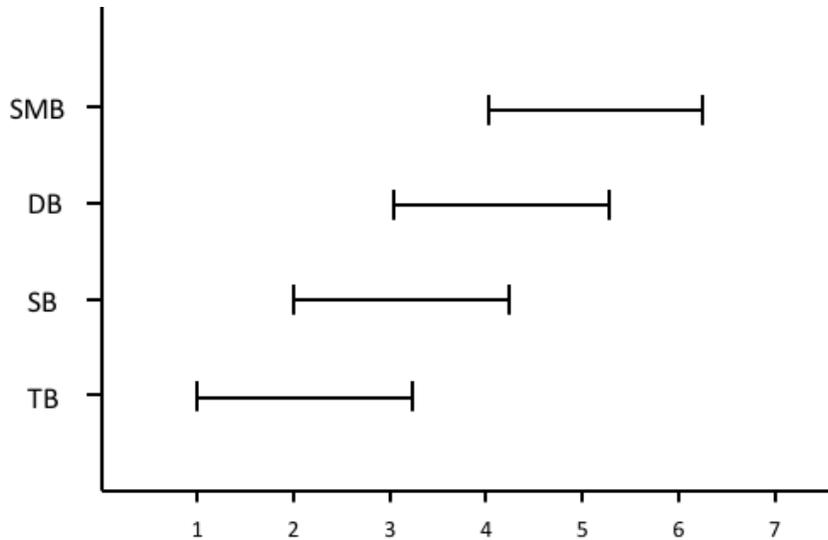


Figure 8.3: Critical difference diagram for the proposed image classification approaches

Dataset	Elsayed (2-D)	Udomchaiporn (3-D)
Epilepsy	<b>88.30</b>	77.30
Musician	<b>99.10</b>	86.10

Table 8.6: Comparison of AUC between the work reported in this thesis (3-D) and previous work (2-D)

[37] were better than those reported in this thesis for both datasets. It is interesting that the best results from both the work in [37] and the work reported in this thesis came from the usage of Tree-based representation techniques. However, it cannot be concluded from the results presented in Table 8.6 that 2-D classification is better than 3-D classification. This is because the corpus callosum might be a better indicator of musical ability and/or the presence of epilepsy than the ventricles. More experiments, using the same parameter settings, but with alternative datasets, related to different application domains, are required before any such conclusion can be made.

## 8.4 Conclusion

In the following chapter, this thesis is concluded with a summary, a review of the main finding with respect to the research objectives presented in Chapter1 and some suggested directions for future work.

## Chapter 9

# Conclusion

This chapter provides a summary of the work presented in this thesis; including: a summary of the proposed approaches to VOIBIC presented in the thesis, the main findings in terms of the research question postulated in the introduction to this chapter and the identified associated research issues, the research contributions and some potential fruitful directions for future research. The chapter is organised as follows. Section 9.1 presents the summary of the work presented in this thesis. The following section, Section 9.2, then presents the main findings in terms of the research question and contribution of this thesis. Ideas for future research directions are then given in Section 9.3.

### 9.1 Summary of the Proposed Approaches

The work presented in this thesis was directed at the classification of 3-D objects found in 3-D data sets, Volume Of Interest Based Image Classification (VOIBIC). To act as a focus for the work MRI brain scan data was used where we wish to classify the images according to the left and right lateral ventricles that are featured in such images. There are four ventricles in a human brain: two lateral ventricles (referred to as the left and right ventricles), a third ventricle which connects to both lateral ventricles and a fourth ventricle that connects the third ventricle with the spinal cord. Only the left and right lateral ventricles were considered with respect to the focus of the presented research. This was because: (i) the lateral ventricles are relatively easy to identify within 3-D MRI brain scans, so facilitating automatic extraction; and (ii) they are much larger than the other two ventricles and consequently can be argued to be more significant.

VOIBIC entails several challenges of which the most significant are: (i) how best to identify (segment) the VOI, and (ii) once identified how these VOI can best be represented to support classification. Although many segmentation algorithms exist these were found to be unsatisfactory, instead two bespoke segmentation algorithms were proposed to achieve the desired segmentation:

1. Volume Growing technique.

2. Bounding Box technique.

The first technique used the volume growing concept to capture the VOI. The process started by manually identifying a starting slice as a guide point whose location was in the ventricle area and then “growing” this point. An acknowledged disadvantage of the Volume Growing algorithm was the need to manually select a start slice and guide point. The second technique comprised three main steps: (i) definition of a bounding box (rectangular shape) that is expected to encompass the VOI (ventricles) with respect to all relevant slices in the given MRI volume, (ii) for each slice collection of the black pixels (voxels), which will include the area of the lateral ventricles, and (iii) application of appropriate noise removal. This process was relatively easy to implement, however, it required bounding boxes to be defined for each image slice. The results of the experiments comparing the operation of the two techniques indicated that the accuracy of the Bounding Box segmentation process was significantly better than the Volume Growing process.

The three VOI representation approaches, to support VOIBIC, considered in this thesis were:

1. Statistical metrics based.
2. Point series based.
3. Tree based.

The first proposed approach was simply founded on the idea of statistical metrics. The idea here was to provide for a benchmark representation to which the two other approaches presented in the thesis could be compared. A number of statistical measures were considered: width, length, height, perimeters in each plane, volume and volume extent of the VOI. These were used to define a ND feature space from which feature vectors for individual VOI could be extracted. For evaluation purposes the statistical feature representation was tested with and without augmentation (the addition of age and gender information). Two selected classification mechanisms, SVM and Decision Tree, were adopted with respect to the evaluation. The reported results indicated that:

1. Reasonable results were obtained, a best accuracy of 78.52% was recorded with respect to the musicians data set and a best accuracy of 86.02% with respect to the epilepsy data set.
2. The augmented data considerably improved the classification results.
3. Overall the approach performed better with respect to the musicians data than the epilepsy data set.

4. Classification effectiveness using Decision Tree classification, on average, was better than when using SVM, but the usage of Decision Trees was more efficient.

The second representation considered in this thesis was founded on the concept of a point series representation coupled with two classification paradigms: (i) feature space classification and (ii) KNN classification. The basic idea was to represent VOI boundaries in terms of a series of points. Two techniques were devised for generating the desired point series: (i) Disc-based and (ii) Spoke-based. For the later a number of different “spoke spacings were considered. For the Feature space classification a feature space was generated using the Hough signature extraction mechanism. For KNN classification the “warping path” distance, obtained using Dynamic Time Warping (DTW) curve comparison, was used as the similarity measure. The evaluation results obtained indicated that:

1. The Spoke-based technique, coupled with KNN classification, generated the better results in terms of both classification effectiveness and efficiency.
2. Spoke-based technique, regardless of the adopted classification paradigm used, out-performed the Disc-based technique in terms of both classification effectiveness and efficiency.
3. A spoke spacing of  $2^\circ$  typically generated the best results in term of classification effectiveness, although a spoke spacing of  $4^\circ$  would clearly be more efficient.

The third representation considered was a tree based representation, specifically an Oct-tree representation. A hierarchical decomposition technique was used to generate the desired Oct-trees, one tree per MRI brain scan. A weighted Frequent Subgraph Mining (FSM) algorithm, gSpan-ATW, was then applied to identify frequently occurring subgraphs (sub-trees) within the Oct-tree representation. The general assumption for the application of the algorithm was that Oct-tree nodes further away from the root were more significant than those nearer the root because the more distant nodes encapsulated a greater level of detail. The identified frequent sub-trees were viewed as defining a feature space which was then used to represent the image set. A given image set could thus be recast into this format so that each image was represented by a feature vector whose elements were some subset of the identified frequent sub-trees making up the feature space. For evaluation purposes, and as in the case of the statistical-metric based representation, two standard classifier generation techniques, SVM and Decision Tree, were used. The reported evaluation results indicated that:

1. The best minimum support threshold ( $\tau$ ) for FSM was between 20% and 30%.
2. The classification effectiveness using SVM was typically better than when using Decision Trees.

3. As before the classification effectiveness for the Musician dataset was better than that for the Epilepsy dataset.

## 9.2 Main Findings and Contributions

In this thesis, the concept of 3-D image classification according to the nature of VOIs contained within 3D image datasets, was considered. Three VOIBIC approaches were proposed. These approaches considered three different VOI representations to maintain the structural information (shape and size) of VOIs coupled with some appropriate classification techniques. The proposed approaches were designed to address the research questions introduced in Chapter 1. In this section, each of the identified research questions is discussed, together with the manner in which the research presented in this thesis addressed each of these individual questions, as follows (research questions from Chapter 1 repeated verbatim).

1. **What is the most appropriate mechanism for identifying VOI, and in particular the left and right ventricle VOI?:**

The Bounding Box segmentation technique, as described in Chapter 4, was found to be the most appropriate technique to identify VOIs. It was able to achieve excellent accuracy in the context of lateral ventricle segmentation.

2. **Assuming that we can establish a process for identifying the VOI, how do we ascertain the quality of this VOI?:**

Although there is no “gold standard” in the case of lateral ventricle shape evaluation “Bland-Altman”, a statistical test for assessing the degree of agreement, indicated excellent quality of the proposed segmentation technique.

3. **Once the VOI have been identified what would be the nature of the VOI representations to be used so as to support classification?:**

The proposed VOIBIC approaches, described in Chapters 5, 6 and 7, were all able to capture the salient elements of the VOI under consideration (the lateral ventricles) so as to retain the structural information (shape and size) of the VOIs. The statistical comparison presented above indicated that the Tree Based representation produced the best results in terms of AUC.

4. **What are the most appropriate parameter settings for the considered representations?:**

Not all the representations considered required parameters. However, the reported evaluation found that with respect to the Spoke-based representation (as proposed in Chapter 6) a Spoke spacing of  $2^\circ$  was best able to capture the required VOI shape information. With respect to the Tree based representation (as

proposed in Chapter 7), the a minimum support threshold value of between 20% and 30% was best able to generate an “optimal” set of frequent subgraphs.

5. **Given a specific representation what is the most appropriate classification model to be used with that representation?:**

It was anticipated that not all the proposed representations would be compatible with a particular classification model. The proposed approaches worked in different manners according to the nature of the proposed VOI representation. The reported evaluation results indicated that: (i) the Statistical metric based representation was most compatible with SVM classification, (ii) the Point series based representation was most compatible with KNN classification and (iii) the Tree based representation with SVM classification. Therefore, as anticipated, it was not possible to identify a best classification model to adopt in the context of VOIBIC.

6. **What is the overall most appropriate VOI classification process in the context of the ventricle application?:**

The reported results indicated that there was a trade-off between classification effectiveness and efficiency. Although the statistical significance analysis presented in Section 8.2 indicated that the Spoke-based Point series approach produced an overall classification effectiveness as good as the Tree based representation, however, the efficiency aspect of the Spoke-based approach was significantly better than the Tree based technique. It is thus possible to conclude that the Spoke-based Point series representation coupled with KNN classification was the most appropriate VOIBIC mechanism in the context of the ventricle application.

7. **Given a VOI representation is there any benefit to be gained by augmenting the data?:**

The evaluation reported earlier on in this thesis indicated that by augmenting the proposed representations with meta data improved classification effectiveness could be achieved. This was regardless of the representation, classification paradigm and dataset used.

Returning to Chapter 1, the overriding research question was: **“Can appropriate volumetric classification techniques that incorporate 3-D spatial relationship information, while at the same time achieving effective performance, be developed given the significant size of volumetric data sets?”**. The work presented in this thesis clearly indicates that the answer to this question is that the proposed VOI representations, coupled with appropriate classification techniques, can clearly classify such images in a way that is both efficient and effective.

The primary contributions of the research work presented in this thesis were listed in Section 1.4 of Chapter 1. For completeness these are listed again here:

1. Two segmentation algorithms, the Volume Growing and Bounding Box algorithms, to extract a particular VOI across an image dataset. Also an image thresholding technique to help partition the images according to foreground and background.
2. An approach to 3-D MRI classification using a representation based on statistical metrics describing the geometry of a detected VOI. More specifically using geometrical features such as: width, length, height, perimeter (with respect to the three primary axes), volume and volume extent, of the VOI.
3. An approach to 3-D MRI classification based on a point series representation coupled with Hough signature extraction. The latter used to produce signatures from the generated point series curves, which were then used as the input to a classification system.
4. An approach to 3-D MRI classification based on a point series representation coupled with the well known K-Nearest Neighbour (KNN) algorithm and Dynamic Time Warping (DTW) to generate a “Warping Distance” as the similarity measure.
5. An approach to 3-D MR classification based on an Oct-tree representation of an identified VOI. The Oct-tree hierarchical representation was coupled with a Frequent Subgraph Mining (FSM) algorithm to identify frequently occurring subtrees (within the Oct-tree representation) which were then used to define a feature space from which feature vectors could be extracted to which an “off the shelf” classification model could be applied.
6. An effective framework for classifying the nature of the left and right ventricles in human MRI brain scans in order to support various kinds of diagnosis and analysis.

### 9.3 Potential Future work

The research described in this thesis has indicated a number of potential research directions for the future. These research directions are briefly introduced in the concluding section of this thesis below.

1. **Alternative 3-D image representation techniques:** A standard feature vector format was adopted with respect to the work described in this thesis, because of its ease of use as it can be used as the input for standard classifiers. It would

be worth while to investigate alternative 3-D representation techniques such as meshes.

2. **Alternative 3-D image classification mechanisms:** A number of established classification models were investigated with respect to the work presented in this thesis. It would be of interest to investigate more classification models to which the proposed representation techniques can be applied.
3. **Deep learning:** Deep learning methods, which are a form of machine learning method, based on learning representations of data could be applied to the work presented in this thesis. Deep learning is based on a set of algorithms that attempt to model high-level abstractions in data by using a “deep graph with multiple processing layers. Hence it could make better representations and create models to learn these representations by simplifying the learning task for large-scale data. Examples of potential deep learning frameworks which can be applied include: deep neural networks, convolutional deep neural networks, deep belief networks and recurrent neural networks.
4. **Alternative domains or datasets:** The evaluation described in this thesis has concentrated on VOIBIC using the lateral ventricles found in 3-D MRI brain scans. A further avenue for investigation would be to investigate the genericness of the proposed approaches by considering other domains or datasets (for example other human organs located in 3-D MRI scans such as the corpus callosum considered in [40, 37]).

Whatever the case, the work presented in this thesis has provided a foundation on which further investigations can be conducted.

# Bibliography

- [1] <http://www.mricro.com>.
- [2] <http://www.mathworks.co.uk/products/matlab>.
- [3] <http://meshlab.sourceforge.net/>.
- [4] Rolf Adams and Leanne Bischof. Seeded Region Growing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(6):641–647, 1994.
- [5] R. Agrawal, R. Srikant, et al. Fast Algorithms for Mining Association Rules. In *Proc. 20th int. conf. very large data bases, VLDB*, volume 1215, pages 487–499, 1994.
- [6] Douglas G. Altman. *Practical Statistics for Medical Research*. Chapman and Hall, 1991.
- [7] L. G. Apostolova, A. E. Green, S. Babakchianian, K. S. Hwang, Y. Chou, A. W. Toga, and P. M. Thompson. Hippocampal Atrophy and Ventricular Enlargement in Normal Aging, Mild Cognitive Impairment and Alzheimer’s Disease. *Alzheimer Disease and Associated Disorders*, 26(1):17, 2012.
- [8] D. Ballard. Generalizing the Hough Transform to Detect Arbitrary Shapes. *Pattern Recognition*, 13(2):111–122, 1981.
- [9] E. Barbu, P. Héroux, S. Adam, and E. Trupin. Clustering Document Images using Graph Summaries. In *Machine Learning and Data Mining in Pattern Recognition*, pages 194–202. Springer, 2005.
- [10] R. Bellazzi and B. Zupan. Predictive Data Mining in Clinical Medicine: Current Issues and Guidelines. *International journal of medical informatics*, 77(2):81–97, 2008.
- [11] R. Bellman and R. Kalaba. On Adaptive Control Processes. *IRE Transactions on Automatic Control*, 4(2), 1959.
- [12] A. Ben-Hur and J. Weston. A Users Guide to Support Vector Machines. In *Data Mining Techniques for the Life Sciences*, pages 223–239. Springer, 2010.

- [13] J. Bondy and U. Murty. *Graph Theory with Applications*, volume 290. Macmillan London, 1976.
- [14] M. Bramer. *Principles of Data Mining*. Springer, 2007.
- [15] M. Brummer, R. Mersereau, R. Eisner, and R. Lewine. Automatic Detection of Brain Contours in MRI Data Sets. *Medical Imaging, IEEE Transactions on*, 12(2):153–166, 1993.
- [16] W. Burger and M. J. Burge. *Digital Image Processing: An algorithmic Introduction Using Java*. Springer, 2008.
- [17] J. Canny. A Computational Approach to Edge Detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 6:679–698, 1986.
- [18] N. G. Cascella, D. J. Schretlen, and A. Sawa. Schizophrenia and Epilepsy: Is There A Shared Susceptibility? *Neuroscience Research*, 63(4):227–235, 2009.
- [19] T. Chow and M. Rahman. A New Image Classification Technique Using Tree-Structured Regional Features. *Neurocomputing*, 70(4):1040–1050, 2007.
- [20] S. Chu, E. Keogh, D. Hart, M. Pazzani, et al. Iterative Deepening Dynamic Time Warping for Time Series. In *SDM*, pages 195–212. SIAM, 2002.
- [21] C. Ciofalo, C. Barillot, and P. Hellier. Combining Fuzzy Logic and Level Set Methods for 3D MRI Brain Segmentation. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, pages 161–164. IEEE, 2004.
- [22] C. Cocosco, A. Zijdenbos, and A. Evans. *Automatic Generation of Training Data for Brain Tissue Classification from MRI*. Springer, 2002.
- [23] C. Cocosco, A. Zijdenbos, and A. Evans. A Fully Automatic and Robust Brain MRI Tissue Classification Method. *Medical image analysis*, 7(4):513–527, 2003.
- [24] D. J. Cook and L. B. Holder. *Mining graph data*. John Wiley & Sons, 2006.
- [25] C Cortes and V Vapnik. Support-Vector Networks. *Machine Learning*, 20(3):273–297, 1995.
- [26] Luciano Da Fontoura Costa and Roberto Marcondes Cesar Jr. *Shape Analysis and Classification: Theory and Practice*. CRC Press, 2001.
- [27] Mark W Craven and Jude W Shavlik. Using Neural Networks for Data Mining. *Future generation computer systems*, 13(2):211–229, 1997.

- [28] K. Dabbs, T. Becker, J. Jones, P. Rutecki, M. Seidenberg, and B. Hermann. Brain Structure and Aging in Chronic Temporal Lobe Epilepsy. *Epilepsia*, 53(6):1033–1043, 2012.
- [29] T. O. Dalaker, R. Zivadinov, D. P. Ramasamy, M. K. Beyer, G. Alves, K. S. Bronnick, O. B. Tysnes, D. Aarsland, and J. P. Larsen. Ventricular Enlargement and Mild Cognitive Impairment in Early Parkinson’s Disease. *Movement Disorders*, 26(2):297–301, 2011.
- [30] N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.
- [31] R. Damadian, M. Goldsmith, and L. Minkoff. NMR in Cancer: XVI. FONAR Image of the Live Human Body. *Physiological chemistry and physics*, 9(1):97–100, 1976.
- [32] J. Demšar. Statistical Comparisons of Classifiers over Multiple Data Sets. *The Journal of Machine Learning Research*, 7:1–30, 2006.
- [33] G. Dougherty. *Digital Image Processing for Medical Applications*. Cambridge Press, 2009.
- [34] M. Dunham. *Data Mining: Introductory and Advanced Topics*. Pearson Education India, 2006.
- [35] R. Edelman and S. Warach. Magnetic Resonance Imaging. *New England Journal of Medicine*, 328(10):708–716, 1993.
- [36] A. Efrat, Q. Fan, and S. Venkatasubramanian. Curve Matching, Time Warping, and Light Fields: New Algorithms for Computing Similarity Between Curves. *Journal of Mathematical Imaging and Vision*, 27(3):203–216, 2007.
- [37] A. Elsayed. Region of Interest-based Image Classification. *PhD Dissertation, Department of Computer Science, University of Liverpool*, November 2011.
- [38] A. Elsayed, F. Coenen, M. García-Fiñana, and V. Sluming. Region of Interest Based Image Categorization. In *International Conference on Data Warehousing and Knowledge Discovery*, pages 239–250. Springer, 2010.
- [39] A. Elsayed, F. Coenen, C. Jiang, M. García-Fiñana, and V. Sluming. Corpus Callosum MR Image Classification. In *Proceedings AI’09*, pages 333–348. Springer, 2009.

- [40] A. Elsayed, F. Coenen, C. Jiang, M. García-Fiñana, and V. Sluming. Corpus Callosum MR Image Classification. *Knowledge-Based Systems*, 23(4):330–336, May 2010.
- [41] A. Elsayed, F. Coenen, C. Jiang, M. García-Fiñana, and V. Sluming. Corpus Callosum MR Image Classification. *Knowledge-Based Systems*, 23(4):330–336, 2010.
- [42] C. Faloutsos and K. Lin. FastMap: A Fast Algorithm for Indexing, Data-mining and Visualization of Traditional and Multimedia Datasets. *SIGMOD Rec.*, 24(2):163–174, May 1995.
- [43] E. Fix and Hodges J. Discriminatory Analysis-Nonparametric Discrimination: Consistency Properties. Technical report, DTIC Document, 1951.
- [44] H. Freeman. On the Encoding of Arbitrary Geometric Configurations. *Electronic Computers, IRE Transactions on*, (2):260–268, 1961.
- [45] M. Friedman. A Comparison of Alternative Tests of Significance For the Problem of M Rankings. *The Annals of Mathematical Statistics*, 11(1):86–92, 1940.
- [46] N. Friedman, D. Geiger, and M. Goldszmidt. Bayesian Network Classifiers. *Machine learning*, 29(2-3):131–163, 1997.
- [47] S. A. Friese, M. Bitzer, D. Freudenstein, K. Voigt, and W. Kker. Classification of Acquired Lesions of The Corpus Callosum with MRI. *Neuroradiology*, 42(11):795–802, 2000.
- [48] M. Galloway. Texture Analysis Using Gray Level Run Lengths. *Computer Graphics and Image Processing*, 4(2):172–179, 1975.
- [49] K. Ganser, H. Dickhaus, R. Metzner, and C. Wirtz. A Deformable Digital Brain Atlas System According to Talairach and Tournoux. *Medical Image Analysis*, 8(1):3–22, 2004.
- [50] D. Gao. Volume Texture Extraction for 3D Seismic Visualization and Interpretation. *Geophysics*, 68(4):1294–1302, 2003.
- [51] R. Gonzalez and R. Woods. Digital Image Processing Reading. *MA: Addison-Wesley*, 1992.
- [52] R. Gonzalez and R. Woods. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.
- [53] F. Gorunescu. *Data Mining: Concepts, Models and Techniques*, volume 12. Springer Science & Business Media, 2011.

- [54] M. Hall, E. Frank, and G. Holmes. The WEKA Data Mining Software: An Update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18, 2009.
- [55] J. Han, M. Kamber, and J. Pei. *Data Mining: Concepts and Techniques*. Elsevier, 2011.
- [56] J. Han, J. Pei, and Y. Yin. Mining Frequent Patterns Without Candidate Generation. In *ACM SIGMOD Record*, volume 29, pages 1–12. ACM, 2000.
- [57] S. Haykin. A Comprehensive Foundation. *Neural Networks*, 2(2004), 2004.
- [58] C. Heil and D. Walnut. Continuous and Discrete Wavelet Transforms. *SIAM review*, 31(4):628–666, 1989.
- [59] G. Hirt, J. Ames, M. Bambach, and R. Kopp. Forming Strategies and Process Modelling for CNC Incremental Sheet Forming. *CIRP Annals-Manufacturing Technology*, 53(1):203–206, 2004.
- [60] S. Ho, L. Bullitt, and G. Gerig. Level-Set Evolution with Region Competition: Automatic 3-D Segmentation of Brain Tumors. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, pages 532–535. IEEE, 2002.
- [61] L. B. Holder, D. J. Cook, S. Djoko, et al. Substructure Discovery in the SUBDUE System. In *KDD workshop*, pages 169–180, 1994.
- [62] P. V. C. Hough. Method and Means for Recognizing Complex Patterns, 1962. US Patent.
- [63] W. Hsu, M. L. Lee, and J. Zhang. Image Mining: Trends and Developments. *Journal of Intelligent Information Systems*, 19(1):7–23, 2002.
- [64] J. Huan, W. Wang, D. Bandyopadhyay, J. Snoeyink, J. Prins, and A. Tropsha. Mining Protein Family Specific Residue Packing Patterns from Protein Structure Graphs. In *Proceedings of the Eighth Annual International Conference on Research in Computational Molecular Biology*, pages 308–315. ACM, 2004.
- [65] J. Huan, W. Wang, and J. Prins. Efficient Mining of Frequent Subgraphs in the Presence of Isomorphism. In *Proceedings of The Third IEEE International Conference on Data Mining*, pages 549–552. IEEE Comput. Soc, 2003.
- [66] A. Inokushi, T. Washio, and H. Motoda. An Apriori-based Algorithm for Mining Frequent Substructure from Graph Data. In *Proceedings of the Forth European Conference on Principles and Practice of Knowledge Discovery in Database*, pages 13–23, 2000.

- [67] F. Itakura. Minimum Prediction Residual Principle Applied to Speech Recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 23(1):67–72, 1975.
- [68] Chris L. Jackins and Steven L. Tanimoto. Oct-trees and Their Use in Representing Three-dimensional Objects. *Computer Graphics and Image Processing*, 14(3):249–270, November 1980.
- [69] D. C. Jackson, W. Irwin, K. Dabbs, J. J. Lin, J. E. Jones, D. A. Hsu, C. E. Stafstrom, M. Seidenberg, and B. P. Hermann. Ventricular Enlargement in New-Onset Pediatric Epilepsies. *Epilepsia*, 52(12):2225–2232, 2011.
- [70] C. Jiang and F. Coenen. Graph-based Image Classification by Weighting Scheme. In *Applications and Innovations in Intelligent System XVI*, pages 63–76. 2009.
- [71] I. Jolliffe. *Principal component analysis*. Wiley Online Library, 2002.
- [72] T. Kapur, W. Grimson, W. Wells, and R. Kikinis. Segmentation of Brain Tissue from Magnetic Resonance Images. *Medical image analysis*, 1(2):109–127, 1996.
- [73] S. Kara and F. Dirgenali. A System to Diagnose Atherosclerosis via Wavelet Transforms, Principal Component Analysis and Artificial Neural Networks. *Expert Systems with Applications*, 32(2):632–640, 2007.
- [74] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. *International journal of computer vision*, 1(4):321–331, 1988.
- [75] H. Khotanlou, O. Colliot, J. Atif, and I. Bloch. 3D brain Tumor Segmentation in MRI using Fuzzy Classification, Symmetry Analysis and Spatially Constrained Deformable Models. *Fuzzy Sets and Systems*, 160(10):1457–1473, 2009.
- [76] M. Kuramochi and G. Karypis. Frequent Subgraph Discovery. In *Proceedings of the IEEE International Conference on Data Mining*, pages 313–320. IEEE Comput. Soc, 2001.
- [77] A. Kusiak, K. H. Kernstine, J. A. Kern, K. A. McLaughlin, and T. L. Tseng. Data Mining: Medical and Engineering Case Studies. In *Industrial Engineering Research Conference*, pages 1–7, 2000.
- [78] S. Lessmann, B. Baesens, C. Mues, and S. Pietsch. Benchmarking Classification Models for Software Defect Prediction: A Proposed Framework and Novel Findings. *Software Engineering, IEEE Transactions on*, 34(4):485–496, 2008.
- [79] J.L. Lin, M. Mula, and B.P. Hermann. Uncovering the Neurobehavioural Comorbidities of Epilepsy over the Lifespan. *The Lancet*, 380(9848):1180–1192, Sep 2012.

- [80] Z. Lin, J. Jin, and H. Talbot. Unseeded Region Growing for 3D Image Segmentation. In *Selected papers from the Pan-Sydney workshop on Visualisation-Volume 2*, pages 31–37. Australian Computer Society, Inc., 2000.
- [81] H. Liu and R. Setiono. Chi2: Feature Selection and Discretization of Numeric Attributes. In *tai*, page 388. IEEE, 1995.
- [82] S. Long and Holder L. B. Graph-based Shape Shape Analysis for MRI Classification. *International Journal of Knowledge Discovery in Bioinformatics*, 2(2):19–33, 2011.
- [83] K. Mardia and T. Hainsworth. A Spatial Thresholding Method for Image Segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 10(6):919–927, 1988.
- [84] B. Modayur, J. Prothero, G. Ojemann, K. Maravilla, and J. Brinkley. Visualization-Based Mapping of Language Function in the Brain. *Neuroimage*, 6(4):245–258, 1997.
- [85] D. Mumford. Mathematical Theories of Shape: Do They Model Perception? In *San Diego, '91, San Diego, CA*, pages 2–10. International Society for Optics and Photonics, 1991.
- [86] W. Niblack. *An Introduction to Digital Image Processing*. Prentice-Hall, 1986.
- [87] R. Niels. Dynamic Time Warping: An Intuitive Way of Handwriting Recognition? 2004.
- [88] V. Niennattrakul and C. Ratanamahatana. Learning DTW Global Constraint for Time Series Classification. *arXiv preprint arXiv:0903.0041*, 2009.
- [89] R. P. Nikhil and K. P. Sankar. A Review on Image Segmentation Techniques. *Pattern Recognition*, 26(9):1277 – 1294, 1993.
- [90] L. Nyúl, A. Falcão, and J. Udupa. Fuzzy-Connected 3D Image Segmentation at Interactive Speeds. *Graphical Models*, 64(5):259–281, 2002.
- [91] W. Oh and B. Lindquist. Image Thresholding by Indicator Kriging. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(7):590–602, 1999.
- [92] V. Ojansivu and J. Heikkilä. Blur Insensitive Texture Classification Using Local Phase Quantization. In *Image and Signal Processing*, pages 236–243. Springer, 2008.
- [93] N. Otsu. A Threshold Selection Method from Gray-Level Histograms. *Automatica*, 11(285-296):23–27, 1975.

- [94] J. Päivärinta, E. Rahtu, and J. Heikkilä. Volume Local Phase Quantization for Blur-Insensitive Dynamic Texture Classification. In *Image Analysis*, pages 360–369. Springer, 2011.
- [95] C. Pham, D. Xu and J. Prince. Current Methods in Medical Image Segmentation 1. *Annual review of biomedical engineering*, 2(1):315–337, 2000.
- [96] J. R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, 1993.
- [97] R. Quinlan. *C4. 5: Programs for Machine Learning*. Elsevier, 2014.
- [98] N. Rajini and R. Bhavani. Classification of MRI Brain Images Using K-Nearest Neighbor and Artificial Neural Network. In *Recent Trends in Information Technology (ICRTIT), 2011 International Conference on*, pages 563–568. IEEE, 2011.
- [99] G.K. Rambally and R. S. Rambally. Octrees and Their Applications in Image Processing. In *Southeastcon'90. Proceedings., IEEE*, pages 1116–1120. IEEE, 1990.
- [100] C. Revol and M. Jourlin. A New Minimum Variance Region Growing Algorithm for Image Segmentation. *Pattern Recognition Letters*, 18(3):249–258, 1997.
- [101] C. Revol-Muller, Fr. Peyrin, Y. Carrillon, and C. Odet. Automated 3-D Region Growing Algorithm Based on an Assessment Function. *Pattern Recognition Letters*, 23(1):137–150, 2002.
- [102] T. Ridler and S. Calvard. Picture Thresholding Using an Iterative Selection Method. *IEEE transactions on Systems, Man and Cybernetics*, 8(8):630–632, 1978.
- [103] N. Robertson and P. Seymour. Graph Minors. II. Algorithmic Aspects of Tree-Width. *Journal of Algorithms*, 7(3):309–322, 1986.
- [104] R. Rojas. *Neural Networks: A Systematic Introduction*. Springer Science & Business Media, 2013.
- [105] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning Representations by Back-Propagating Errors. *Cognitive modeling*, 5(3):1, 1988.
- [106] H. Sakoe and S. Chiba. Dynamic Programming Algorithm Optimization for Spoken Word Recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 26(1):43–49, Feb 1978.
- [107] H. Samet. The Quadtree and Related Hierarchical Data Structures. *ACM Computing Surveys (CSUR)*, 16(2):187–260, 1984.

- [108] J. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*, volume 3. Cambridge university press, 1999.
- [109] L. G. Shapiro and G. C. Stockman. *Computer Vision*. Pearson, 2001.
- [110] 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.
- [111] J. Sijbers, P. Scheunders, M. Verhoye, A. Van der Linden, D. Van Dyck, and E. Raman. Watershed-based Segmentation of 3D MR Data for Volume Quantization. *Magnetic Resonance Imaging*, 15(6):679–688, 1997.
- [112] P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer Science & Business Media, 2013.
- [113] J. Soni, U. Ansari, D. Sharma, and S. Soni. Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction. *International Journal of Computer Applications*, 17(8):43–48, 2011.
- [114] T. L. Stedman. *Stedman’s Medical Dictionary for the Health Professions and Nursing*. Lippincott Williams and Wilkins, 2012.
- [115] P. Striano, M. M. Mancardi, R. Biancheri, F. Madia, E. Gennaro, R. Paravidino, F. Beccaria, G. Capovilla, B. D. Bernardina, F. Darra, et al. Brain MRI Findings in Severe Myoclonic Epilepsy in Infancy and Genotype–Phenotype Correlations. *Epilepsia*, 48(6):1092–1096, 2007.
- [116] H. Suzuki and J. Toriwaki. Automatic Segmentation of Head MRI Images by Knowledge Guided Thresholding. *Computerized medical imaging and graphics*, 15(4):233–240, 1991.
- [117] Xiaou Tang. Texture Information in Run-Length Matrices. *Image Processing, IEEE Transactions on*, 7(11):1602–1609, 1998.
- [118] F. Tao, F. Murtagh, and M. Farid. Weighted Association Rule Mining Using Weighted Support and Significance Framework. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 661–666. ACM, 2003.
- [119] S. Theodoridis and K. Koutroumbas. *Pattern Recognition*. 2003.
- [120] D. Thom. Dilatation of the Lateral Ventricles as a Common Brain Lesion in Epilepsy. *The Journal of Nervous and Mental Disease*, 46(5):355–358, 1917.

- [121] B. Trier and A. Jain. Goal-Directed Evaluation of Binarization Methods. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(12):1191–1201, 1995.
- [122] A. Udomchaiporn, F. Coenen, M. García-Fiñana, and V. Sluming. 3-D MRI Brain Scan Feature Classification Using an Oct-Tree Representation. *Advanced Data Mining and Applications*, 8346:229–240, 2013.
- [123] A. Udomchaiporn, F. Coenen, M. García-Fiñana, and V. Sluming. 3-D MRI Brain Scan Classification of Epilepsy Versus Non-epilepsy. *Medical Image Understanding and Analysis 2014*, pages 253–258, 2014.
- [124] A. Udomchaiporn, F. Coenen, M. García-Fiñana, and V. Sluming. 3-D MRI Brain Scan Classification Using A Point Series Based Representation. *Data Warehousing and Knowledge Discovery*, 8646:300–307, 2014.
- [125] A. Udomchaiporn, F. Coenen, M. García-Fiñana, and V. Sluming. 3-D Volume of Interest Based Image Classification. *The Pacific Rim International Conference on Artificial Intelligence*, 2016.
- [126] J. Udupa and S. Samarasekera. Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation. *Graphical models and image processing*, 58(3):246–261, 1996.
- [127] J. Vial, H. Noçairi, P. Sassiati, S. Mallipatu, G. Cognon, B. Thiébaud, D. and Teillet, and D. N. Rutledge. Combination of Dynamic Time Warping and Multivariate Analysis for the Comparison of Comprehensive Two-Dimensional Gas Chromatograms: Application to Plant Extracts. *Journal of Chromatography A*, 1216(14):2866–2872, 2009.
- [128] L. Vincent. Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms. *Image Processing, IEEE Transactions on*, 2(2):176–201, 1993.
- [129] L. Vincent and P. Soille. Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (6):583–598, 1991.
- [130] M. Vlachos, Z. Vagena, P. S. Yu, and V. Athitsos. Rotation Invariant Indexing of Shapes and Line Drawings. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management, CIKM’05*, pages 131–138, 2005.

- [131] X. Wang, T. X Han, and S. Yan. An HOG-LBP Human Detector with Partial Occlusion Handling. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 32–39. IEEE, 2009.
- [132] O. Wirjadi. *Survey of 3d Image Segmentation Methods*. ITWM, 2007.
- [133] S. F. Witelson. Hand and Sex Differences in the Isthmus and Genu of the Human Corpus Callosum. *Brain*, 112(3):799–835, 1989.
- [134] X. Wu, V. Kumar, R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. McLachlan, A. Ng, B. Liu, Y. Philip, et al. Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1):1–37, 2008.
- [135] X. Yan, S. Yu, and J. Han. Graph Indexing: A Frequent Structure-Based Approach. In *Proceedings of the 2004 ACM SIGMOD international conference on Management of data*, pages 335–346. ACM, 2004.
- [136] Xifeng Yan. gSpan: Graph-based Substructure Pattern Mining. In *Proceeding of The IEEE International Conference on Data Mining*, pages 721–724. IEEE Comput. Soc, 2002.
- [137] D. Yu, X. Yu, Q. Hu, J. Liu, and A. Wu. Dynamic Time Warping Constraint Learning for Large Margin Nearest Neighbor Classification. *Information Sciences*, 181(13):2787–2796, 2011.
- [138] J. Yu, J. Amores, N. Sebe, P. Radeva, and Q. Tian. Distance Learning for Similarity Estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(3):451–462, 2008.
- [139] Y. Zhang. A Survey on Evaluation Methods for Image Segmentation. *Pattern recognition*, 29(8):1335–1346, 1996.
- [140] Y. Zhang, Z. Dong, L. Wu, and S. Wang. A Hybrid Method for {MRI} Brain Image Classification. *Expert Systems with Applications*, 38(8):10049 – 10053, 2011.
- [141] G. Zhao and M. Pietikainen. Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(6):915–928, 2007.