IMPROVED FEATURE EXTRACTION AND CLASSIFICATION TECHNIQUES FOR MULTISPECTRAL BRAIN MAGNETIC RESONANCE IMAGES

Ph D Thesis submitted to Cochin University of Science and Technology in partial fulfillment of the requirements for the award of the Degree of

DOCTOR OF PHILOSOPHY

By

SINDHUMOL S
(Reg.No. 3764)

Under the guidance of

Dr. B KANNAN

Department of Computer Applications
Cochin University of Science and Technology
Cochin -6820 22, Kerala, India

July 2013

Improved Feature Extraction and Classification Techniques for Multispectral Brain Magnetic Resonance Images

Ph D thesis

Author:

Sindhumol S
Research Scholar
Department of Computer Applications
Cochin University of Science & Technology
Cochin -682022, Kerala, India
sindhumol_s@cusat.ac.in, sindhumol09@gmail.com

Research Advisor:

Dr. B. Kannan
Associate Professor
Department of Computer Applications
Cochin University of Science & Technology
Cochin -682022, Kerala, India
bkannan@cusat.ac.in

Department of Computer Applications
Cochin University of Science & Technology
Cochin -682022, Kerala, India
July 2013

Cover image: MRI

Dr B Kannan (Supervising Guide)
Associate Professor
Dept of Computer Applications
Cochin University of Science and Technology

Certificate

*Certified that the work presented in this thesis entitled "**Improved Feature Extraction and Classification Techniques for Multispectral Brain Magnetic Resonance Images**" is based on the bonafide research work done by **Sindhumol S** under my guidance in the Department of Computer Applications, Cochin University of Science and Technology, Kochi-22 and has not been included in any other thesis submitted previously for the award of any degree.*

Kochi-22

03-07-2013 ***Dr B Kannan***

Dr B Kannan (Supervising Guide)
Associate Professor
Dept of Computer Applications
Cochin University of Science and Technology

Certificate

This is to certify that all the relevant corrections and modifications suggested by the audience during the Pre-synopsis seminar and recommended by the Doctoral Committee of the candidate have been incorporated in the thesis.

Kochi-22

03-07-2013 *Dr B Kannan*

DECLARATION

I hereby declare that the present work entitled ***"Improved Feature Extraction and Classification Techniques for Multispectral Brain Magnetic Resonance Images"*** *is based on the original work done by me under the guidance of Dr B Kannan, Associate Professor, Department of Computer Applications, Cochin University of Science and Technology, Kochi-22 and has not been included in any other thesis submitted previously for the award of any other degree.*

Kochi-22 ***Sindhumol S***

03-7-2013

Acknowledgements

I would like to thank **God Almighty** in the first place, who gave the determination and courage to overcome many trying moments to pursue my dreams.

Next, my sincere and special gratitude goes to my supervisor, **Dr. B. Kannan**, who has accepted me as a PhD student and provided me with his invaluable academic guidance and support during this research. His energy, enthusiasm, and inspirations always encouraged me to elevate and promote my work to a level that would not have been otherwise possible.

Besides my advisor, I am grateful to **Dr. Pramod K.V**. of my department for all his support throughout my research period.

I take this opportunity to thank **Dr. A. Sreekumar** of the same department for his inspiration and advice during these years.

This work would never have been accomplished without generous help and encouragement provided by medical expert, **Dr. Anil Kumar**, Senior Consultant Radiologist at IRIS, Medall Centre in Kochi. All the administrative staff and Technicians of IRIS were so co-operative and helpful every time I visit IRIS for data collection.

My sincere acknowledgement goes to **Dr. C. Kesavadas** (Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananathauram) for his inspiration and support to frame out my carrier in MRI analysis.

In particular, I wish to express my gratitude to **Prof. Kamal Bijlani**, and all **my friends** in Amrita E-Learning Research Lab for their continuous support and encouragement to fulfill my Ph.D. I feel very fortunate and happy to have such a great workplace and friends in my life.

I whole-heartedly thank my friend **Dr. Jeny Rajan**, who helped me a lot to collect journals and research materials. His passionate and persistent attitude

towards the experimental research motivated me in my work and helped me to progress through many bottlenecks.

I also thank all non-teaching staff of Department of Computer Application, CUSAT for all their timely help and immense support in the completion of my work.

I thank all the research scholars of my department especially **Simily Joseph, Jomy John, Ramkumar, Bino, Binu, Tibin, and Remya** for their valuable ideas and suggestions throughout my research period.

I would like to extend my gratitude to my teachers **Prof. Dr. M. R. Kaimal,** and **Prof. Dr. M. Wilscy**, whose lectures and support inspired me to start my research studies in medical image analysis.

Finally—and most importantly—I would like to thank my family for their love, patience, and constant eagerness to help in any way that they can. Without their encouragement and understanding it would have been impossible for me to finish this work. I would like to dedicate this thesis to **my parents** who always wanted me to get the best. I am sure my father is now smiling at me from heaven.

Sindhumol S

Contents

List of Figures

List of Tables

List of Abbreviations

AC	Approximation Coefficient
ANN	Artificial Neural Networks
AUC	Area Under the Curve
BEP	Band-Expansion Process
BSS	Blind Source Separation
CAD	Computer-Aided Diagnosis
CLT	Central Limit Theorem
CSF	CerebroSpinal Fluid
CV	Cross Validation
DC	Detail Coefficient
DWT	Discrete Wavelet Transform
ECV10	External 10-fold CV
EM	Expectation Maximization
FCM	Fuzzy C-Means
FLAIR	Fluid-Attenuated Inversion Recovery
fMRI	functional MRI
FNR	False Negative Rate
FPR	False Positive Rate
GM	Gray Matter
ICA	Independent Component Analysis
ICA+SVM	ICA based SVM
KICA	Kernel Independent Component Analysis
k-NN	k-Nearest Neighbour
MICA+SVM	MICA based SVM
ML	Maximum Likelihood
MRI	Magnetic Resonance Imaging
MR	Magnetic Resonance

MS	Multiple Sclerosis
MW-ICA	Multisignal Wavelet ICA
OC-ICA	Over Complete-ICA
PCA	Principal Component Analysis
PD	Proton Density
PNN	Probabilistic Neural Networks
RF	Radio Frequency
ROC	Receiver Operating Characteristic
SAM	Spectral Angle Mapping
SC-ICA	Spectral Clustering ICA
SD	Standard Deviation
SOM	Self Organizing Map
SPM	Statistical Parametric Mapping
SVM	Support Vector Machines
T1W	T1-Weighted
T1WI	T1-Weighted Image
T2W	T2-Weighted
T2WI	T2-Weighted Image
TE	Echo Time
TI	Tanimoto Index
TPR	True Positive Rate
TR	Repetition Time
WM	White Matter
WML	White Matter Lesion

Abstract

Magnetic Resonance Imaging (MRI) is a multi sequence medical imaging technique in which stacks of images are acquired with different tissue contrasts. Simultaneous observation and quantitative analysis of normal brain tissues and small abnormalities from these large numbers of different sequences is a great challenge in clinical applications. Multispectral MRI analysis can simplify the job considerably by combining unlimited number of available co-registered sequences in a single suite. However, poor performance of the multispectral system with conventional image classification and segmentation methods makes it inappropriate for clinical analysis. Recent works in multispectral brain MRI analysis attempted to resolve this issue by improved feature extraction approaches, such as transform based methods, fuzzy approaches, algebraic techniques and so forth. Transform based feature extraction methods like Independent Component Analysis (ICA) and its extensions have been effectively used in recent studies to improve the performance of multispectral brain MRI analysis. However, these global transforms were found to be inefficient and inconsistent in identifying less frequently occurred features like small lesions, from large amount of MR data.

The present thesis focuses on the improvement in ICA based feature extraction techniques to enhance the performance of multispectral brain MRI analysis. Methods using spectral clustering and wavelet transforms are proposed to resolve the inefficiency of ICA in identifying small abnormalities, and problems due to ICA over-completeness. Effectiveness of the new methods in brain tissue classification and segmentation is confirmed by a detailed quantitative and qualitative analysis with synthetic and clinical, normal and abnormal, data. In comparison to conventional classification techniques, proposed algorithms provide better performance in classification of normal brain tissues and significant small abnormalities.

List of Publications

- Sindhumol S., Anilkumar, Kannan Balakrishnan,*"Spectral clustering independent component analysis for tissue classification from brain MRI"*, **Bio-medical Signal processing and Control, Elsevier,** Vol. 8, Issue 6, November 2013, pp. 667–674. http://dx.doi.org/10.1016/j.bspc.2013.06.007

- Sindhumol S., Anilkumar, Kannan Balakrishnan, *"Abnormality Detection from Multispectral Brain MRI using Multiresolution Independent Component Analysis"*, **International Journal of Signal Processing, Image Processing and Pattern Recognition**, Vol. 6, No. 1, February, 2013, pp. 177-190.

- Sindhumol S., Anilkumar, Kannan Balakrishnan,*"Automated Brain Tissue Classification by Multisignal Wavelet Decomposition and Independent Component Analysis"*, **ISRN Biomedical Imaging**, Volume 2013, Article ID 473437. http://dx.doi.org/10.1155/ 2013/473437

- Sindhumol S, Kannan Balakrishnan, Anil Kumar, "Brain Tissue Classification from Multispectral MRI by Wavelet based Principal Component Analysis", **IJIGSP**, vol.5, no.8, pp.29-36, 2013. http://dx.doi.org/10.5815/ijigsp.2013.08.04 .

- Sindhumol S., Anilkumar, Kannan Balakrishnan, *"Wavelet based Independent Component Analysis for Multispectral Brain Tissue Classification"*, **Proc. IEEE International Conference on Communication and Signal Processing - ICCSP' 13**, April 2013, India. http://dx.doi.org/10.1109/iccsp.2013.6577086

- Sindhumol S., Anilkumar, Kannan Balakrishnan,*"A hybrid approach to brain tissue classification using support vector machines"*, **Applied Soft Computing, Elsevier** (Submitted).

Chapter -1
INTRODUCTION

Magnetic Resonance Imaging (MRI) is an image acquisition technique in radiology to visualize the structure and function of the body. Unlike Computed Tomography (CT) scanning, MRI uses no ionizing radiation, and is generally a very safe procedure. This multi- sequence digital imaging technique acquires stack of images with different tissue contrasts, for example T1-Weighted (T1W), T2-Weighted (T2W), Proton Density (PD) and so forth [1]. Each sequence highlights specific properties of tissues and pathologies, which is useful in diagnosis of neurological (brain), musculoskeletal, cardiovascular, and oncological (cancer) diseases [2, 3]. However, simultaneous observation and quantitative analysis of small details and abnormalities from these large numbers of different sequences is a great challenge in clinical applications. Multispectral MRI analysis [1, 4] can simplify the job considerably by combining unlimited number of available co-registered sequences in a single suite. However, conventional algorithms used in normal data mining process are not efficient and robust to provide the required clinical accuracy in brain tissue classification. Recent multispectral brain MRI analysis attempted to resolve this issue by improved feature extraction with transform based methods [5], fuzzy approaches [6], algebraic techniques and so forth [7, 8].

This dissertation describes the effectiveness of transform based feature extraction techniques, Independent Component Analysis (ICA) and multisignal wavelet analysis, in multispectral classification of Magnetic Resonance (MR) brain images. This chapter gives an introduction of the

nature and significance of the problem domain, and outlines the motivations and contributions.

1.1 Multispectral analysis of brain MRI

1.1.1 Magnetic Resonance Imaging (MRI)

In MRI, the principles of Nuclear Magnetic Resonance (NMR) are used to obtain the microscopic, chemical and physical information about molecules [2]. Nobel Laureate Professor Sir Peter Mansfield from University of Nottingham and Nobel Laureate Professor Paul Christian Lauterbur from the University of Illinois made NMR capable of producing images of biological body, and in 1973, they demonstrated MRI using the back projection ideas introduced in CT [9]. The basic mechanism behind MR scanner is the approach of aligning magnetization vectors of hydrogen atoms by application of an external magnetic field, as shown in Fig. 1.1 [10]. An MR scanner contains strong magnets, arranged in the circular part of the scanner. The patient lies flat in the scanner bed and desired part of their body is examined by sliding them into the scanner. Most of the human body is made up of water molecules, which consist of hydrogen and oxygen atoms. A smaller particle, called proton, exists at the centre of each hydrogen atom. They are very sensitive to magnetic fields, and the magnetic moments of the individual hydrogen nuclei are oriented in random directions. When these nuclei are caught suddenly in a strong magnetic field, they lines up in the direction of the applied magnetic field like so many compass needles aligning with the earth's magnetic field. Short bursts of radio waves, RF pulses, are sent to certain areas of the body and the protons are pulled out of position. As this happens, each proton transmits a radio signal that provides information about its exact location in the body.

The radio signals from millions of protons can be collected together and combined to generate a detailed image of different body parts, with the help of a concept, K-Space [11]. K-Space is a matrix of voxels that stores the raw imaging data from the MR imaging system [12].

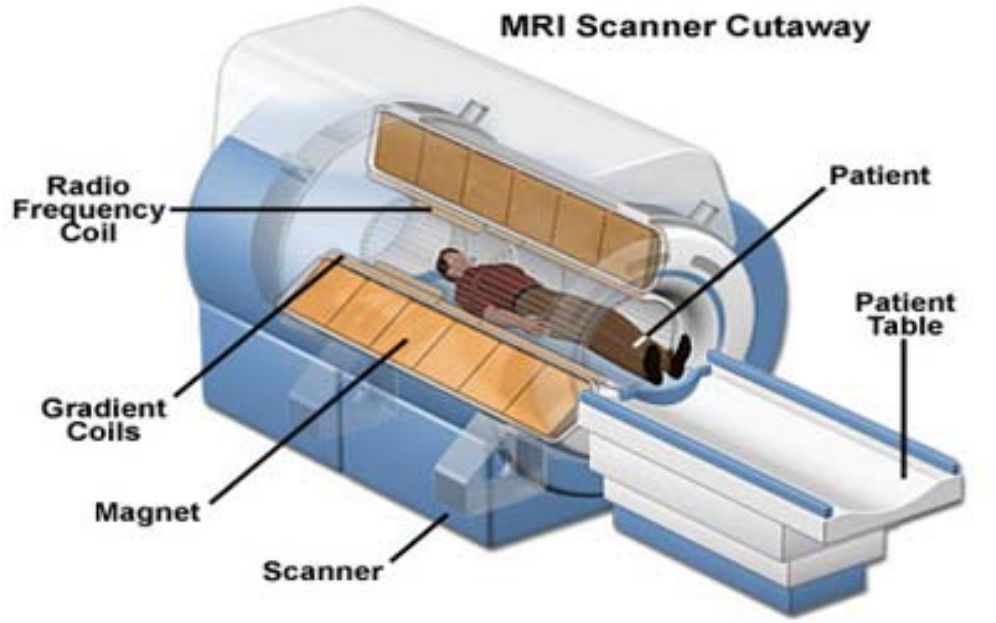


Fig. 1.1 MRI acquisition

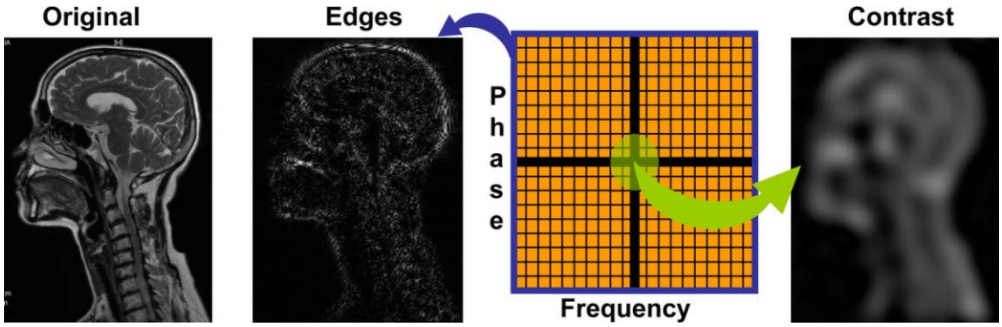


Fig. 1.2 K-space [13]

Usually, the horizontal axis (x-axis) of the matrix corresponds to the frequency, and the vertical axis (y-axis) corresponds to the phase as given in

Fig. 1.2 [13]. Axes of frequency and phase can be easily interchanged also. i.e, K-space data contains all imaging information in its amplitude and phase. Contrast information (low spatial frequency information) is contained in the central part of k-space, whereas details and fine structures (high frequency information) are contained in the periphery of k-space. The raw imaging data in k-space must be Fourier transformed to obtain the final image [13].

The basic idea of utilizing the water molecule in imaging makes MRI most desirable in disease detection, since majority of the diseases manifest themselves by an increase in water content. However, a contrast based assessment of pathology is sometimes difficult; for example, infection and tumor looks similar in some cases. The correct diagnosis can only be done by a careful image analysis results from an experienced radiologist.

1.1.2 Characteristics of brain MRI data

In MR image acquisition, the relaxation process in resonance is controlled by the tissue dependent parameters T1 and T2. They provide means of differentiating among different tissues. T2 is the time constant, characterizing the rate at which excited nuclei exchange energy, or lose phase coherence with each other. It is referred to as the spin-spin (or transverse) relaxation time because it is the loss of transverse magnetization that determines the T2 relaxation time [14]. T1 is the time constant characterizing the rate at which excited nuclei dissipate excess energy to the environment (lattice). It is referred to as the spin-lattice relaxation time (or longitudinal relaxation time).

1.1.2.1 Repetition Time and Echo Time

Repetition Time (TR) is the time (in milliseconds) between the application of an RF excitation pulse and the start of the next RF pulse [15, 16]. TE (also measured in milliseconds) is the time between the application of the RF pulse and the peak of the echo detected [15, 16]. Both parameters affect contrast on

MR images because they provide varying levels of sensitivity to differences in relaxation time between various tissues. At short TRs, the difference in relaxation time between fat and water can be detected (longitudinal magnetization recovers more quickly in fat than in water). However, it is not possible for long TRs. At short TEs, differences in the T2 signal decay in fat and water cannot be detected; but at long TEs, they can be detected. Therefore, TE relates to T2 and affects contrast on T2-weighted images.

1.1.2.2 Tissue contrast

Tissue contrast in all MR images are affected by each of the parameters, T1, T2, PD etc. to some degree. Also, TR and TE can be adjusted to emphasize a particular type of contrast. In T1-weighted MR imaging (Fig. 1.3), while the images show all types of contrast, T1 contrast is made more prominent.

T1W images best describe the anatomy, and, if contrast material is used, they also may show pathologic entities; however, T2W images (Fig. 1.3) provide the best details of disease, because most tissues that are involved in a pathologic process have a higher water content than is normal, and the fluid causes the affected areas to appear bright on T2W images [15]. PD images (Fig. 1.3) usually show both the anatomy and the disease entity.

Latest MR scans often consist of another variation, known as Fluid-Attenuated Inversion Recovery (FLAIR) (Fig. 1.3). It is a heavily T2W imaging technique that suppresses the ventricular CerebroSpinal Fluid (CSF) signal. The strongest signals are usually collected from certain brain abnormalities. Therefore, CSF and abnormalities appear black and bright, respectively. Although FLAIR is similar to T2W imaging, it allows better abnormality visualization since signal of free water is suppressed [17].

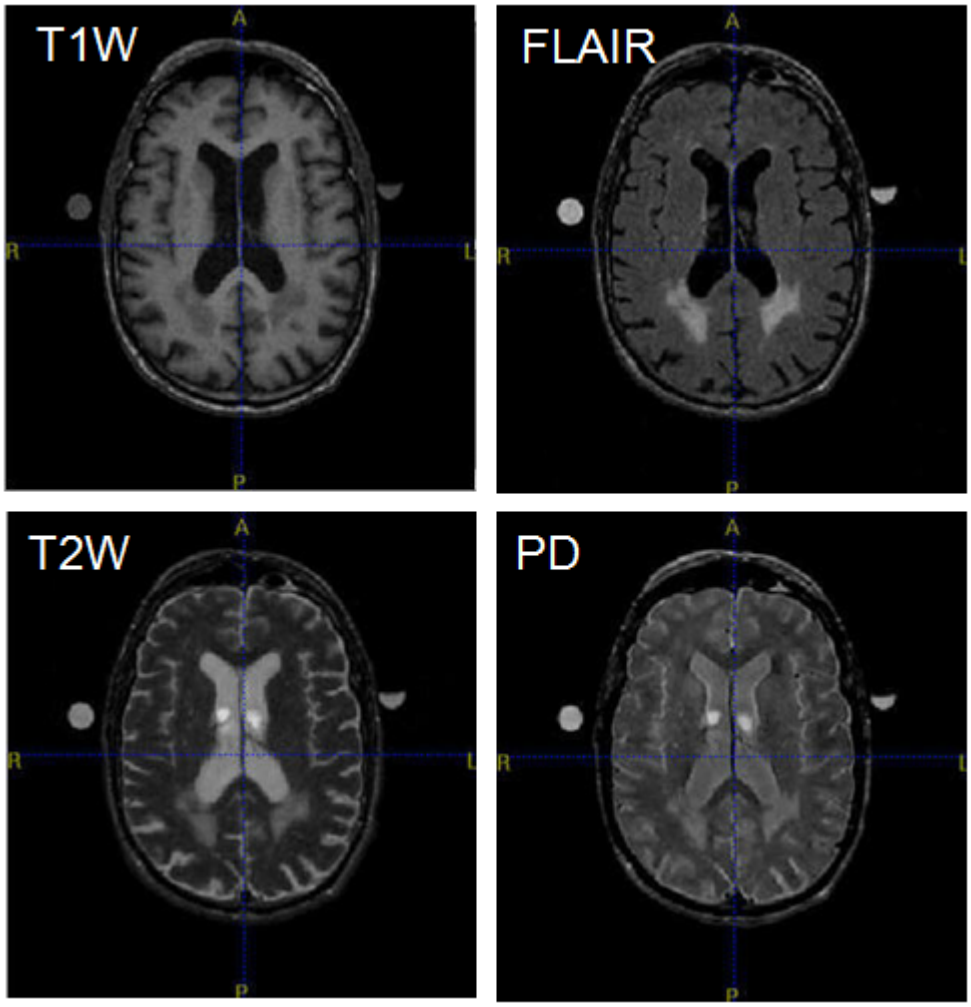


Fig. 1.3 MRI modalities [18]

These MRI modalities are often combined together to facilitate a more accurate analysis, referred to as multi-spectral image analysis [3, 7], where more than one measurement is made at each location in the image [19]. Techniques in multi-spectral MRI offer medical practitioners more information to characterize and discriminate various tissues based on physical and biochemical properties [4, 7].

In this thesis work, we considered both synthetic and clinical data for multispectral analysis. Synthetic MR images contains normal and abnormal

(multiple sclerosis) data obtained from BrainWeb[20], the Simulated Brain Database at the McConnell Brain Imaging Centre of the Montreal Neurological Institute (MNI), McGill University. Axial T1-Weighted Images (T1WI), T2-Weighted Images (T2WI), and PD Images (PDI) formed the input multispectral suite. Slices from each sequence have parameter settings, 1-mm slice thickness, 0% intensity non-uniformity and noise level, 0%.

Table 1.1 Clinical input data characteristics

Modalities	TR (ms)	TE (ms)	IR (ms)	Slice thickness (mm)	Slice Gap (mm)	Size
T1W	1600	8.9		5	6.5	320x320
T2W	4000	95				344x384 416x512
FLAIR	6000 9000	94	2026.5 2500			464x512

TR – Repetition time, TE –Echo Time, IR- Inversion Recovery time,

Total 398 healthy and unhealthy cases sampled by Siemens' whole body 3T MR system (Siemens, AG Medical Solutions, Erlangen, Germany) were considered as a part of data collection. From these, we avoided 16 cases due to severe artifacts. Selected axial T1WI, T2WI and FLAIR images were of characteristics as given in Table 1.1. Axial spin echo T1WI with repetition time (TR) = 1600ms, echo time (TE) = 8.9ms, image dimension 320x320 pixels, and T2WI with TR/TE = 4000ms/95ms. Both 384 mm and 512 mm resolutions were considered for T2W, with sizes 344x384 and 416x512 pixels respectively. FLAIR images with image dimension 464x512 pixels, TR/TE/ inversion time = 6000ms/94ms/2026.5ms for abnormal cases and with TR/TE/inversion time = 9000ms/94ms/2500ms for normal cases were also included in the analysis.

1.1.3 Multispectral image analysis

Multispectral approach was originally developed for applications using satellite images, to analyze multispectral data for NASA's LANDSAT series [21]. First attempt in brain MRI analysis was made by Vannier *et al.* (1985) [22], and second work was with breast MRI [23]. Pixel intensity based studies using more than three MR channel images in the multispectral analysis were introduced in 1994 by Taxt and Arvid Lundervold [24], after which it extended its applications to different areas of MRI analysis.

Pattern recognition techniques are generally considered as the most effective methods for multi-spectral image analysis, where the classification methods are divided into unsupervised and supervised learning. Unsupervised methods like Expectation Maximization (EM), k-means [4] and its fuzzy equivalent, the most widely used Fuzzy C-Means (FCM) [7] generally creates satisfactory results in MR image analysis [24,25]. However, clustering is not a reliable method for accurate classification in pathological analysis [24]. The conventional supervised learning machines like Artificial Neural Networks (ANN) [26], Probabilistic Neural Networks (PNN) [27], and Support Vector Machines (SVM) [28] have been effectively used in multispectral MRI analysis. However, application of these conventional classification methods alone in multispectral MRI analysis often failed to provide expected clinical accuracy [7, 29].

A typical multispectral analysis system initializes with a collection of co-registered images as input data. Corresponding slices from each sequence, T1W, T2W, PD etc are considered as each band in the multispectral suite as shown in Fig. 1.4.

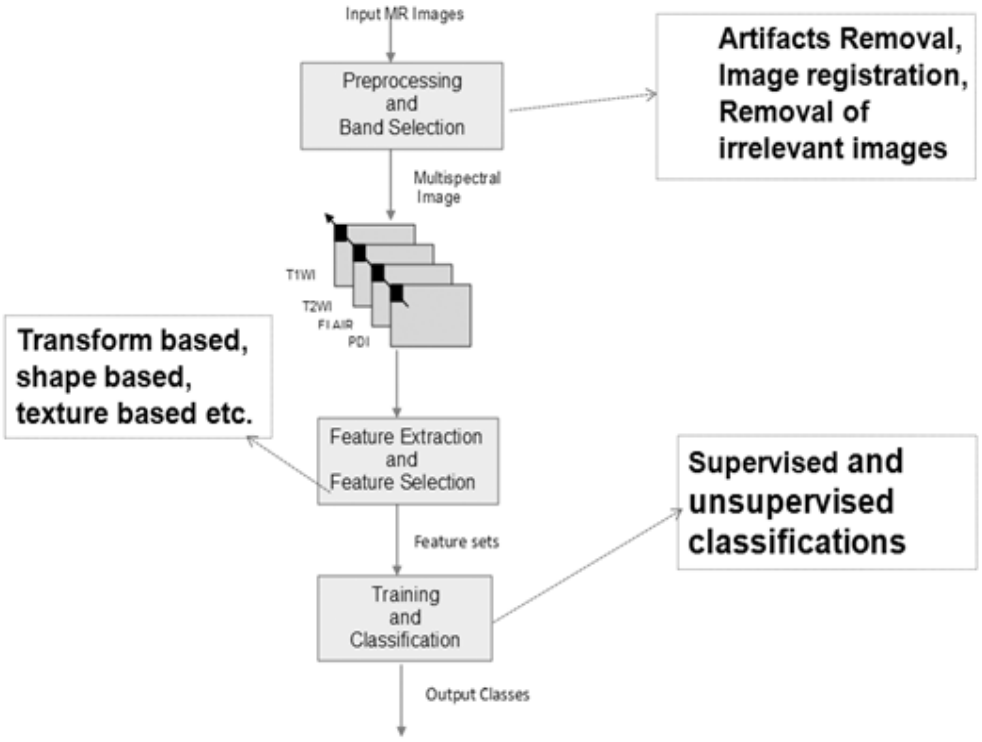


Fig. 1.4 Typical multispectral analysis

Preprocessing steps like intensity inhomogeneity correction and noise removal [7] can be used to enhance the contrast and quality of the input images. Feature extraction, and optimal selection of features is important in multispectral analysis, since accuracy and reliability of the classified results are highly affected by the selected feature sets. Recent researches on feature extraction using transform based methods such as Principal Component Analysis (PCA) [30, 31], ICA [5], and wavelets [32] have contributed a lot in high performance brain tissue classification and diagnosis. The present thesis focuses on design, implementation, and evaluation of new ICA extensions using wavelets and spectral clustering for multispectral MRI analysis.

1.2 Motivation of the work

MRI is a powerful visualization technique in clinical practice and biomedical research for investigation of brain anatomy and function. It provides much greater contrast between different soft tissues of the body than CT images, which makes it particularly useful in clinical diagnosis, especially in evaluating brain tumors and lesions. Brain matter segmentation and classification from MR sequences is an important image processing step for both medical practitioners and scientific researchers in pathological analysis. Computer Aided Diagnosis (CAD) systems help them to assess the tumor growth and treatment responses. In addition to this, it can assist them in computer-aided surgery, radiation therapy and for modeling the tumor growth. Large amounts of research efforts have been made in developing effective segmentation and classification methods in the past years [7, 24, 25]. However, such methods failed to reach the accuracy level provided by the visual analysis from human experts.

The fast evolution of MR imaging techniques offers a wide repository of pulse sequences that can easily be tuned to offer specific visualizations of the brain. These sequences have high spatial resolution and provide much information on the anatomical structure, allowing quantitative pathological or clinical studies. For example, some lesions are obvious in FLAIR, but its presence should be confirmed in other sequences (T2W or PD) also to avoid false positives. T1W images also can give very useful information to provide an improved segmentation. Slice by slice examination of these sequences is a tedious job in clinical analysis. Therefore, neuro-radiologists demand a new approach of computer-aided diagnosis due to their heavy workloads in extraction of relevant information from large amount of data.

Multispectral image analysis was proposed in the literature [1, 4] as a solution to this, to improve the accuracy and reduce the analysis overhead. However, it was observed that conventional classification methods failed to provide high performance with such a large amount of data. Therefore, efficient feature extraction methods are necessary to extract the significant tissue details simultaneously from a large amount of unique information, inherent in different MRI pulse sequences.

Multivariate analysis method, ICA, was found to be effectively used in feature extraction from multispectral brain MRI studies [33]. ICA can efficiently unmix the brain substances into different Independent Components (ICs). Recent attempts to resolve the limitations of ICA, over-completeness and random initial projection vectors, in MRI analysis improved the brain tissue analysis to a great extent [5, 29]. However, another major challenge still exist; the loss of less frequently occurred features such as small lesions, while processing a large amount of MRI data. Therefore, in addition to methods resolving over-completeness and inconsistency in ICA, other object based local feature extraction techniques to retain the relevance of small abnormalities are desirable in multispectral brain tissue analysis.

In an attempt to achieve this, an idea of retaining the local features through a pre-processing technique or feature extraction technique is investigated in this thesis. In order to implement this idea, spectral clustering approaches and wavelet based methods are introduced along with ICA, and evaluated with supervised and unsupervised classification methods.

1.3 Objectives of the thesis

The primary objective of this research is to design and develop methods and algorithms to improve the efficiency of brain MR image analysis. Specifically, this thesis work focus on design of efficient, transform based feature extraction methods, to improve the simultaneous brain tissue classification and pathological study from multiple MR sequences. To achieve this, the current research work has concentrated on the following goals,

- Identify the major challenging issues in brain MRI analysis, and solve some of these issues by closing the gap between medical and technical disciplines.
- Design and implement a multispectral analysis system to perform simultaneous brain tissue classification and pathological study from multiple MR sequences.
- Improve the performance of brain tissue analysis by providing sophisticated feature extraction techniques.
- Apply the enhanced feature extraction methods to investigate the improvement in extraction of small details from dominating background and other tissues, and evaluate the overall accuracy and efficiency of the multispectral MRI analysis system.

1.4 Contributions

1.4.1 Technical contributions

- Extended Independent Component analysis (ICA) techniques are designed and implemented for improved feature (both global and local) extraction from multispectral MR images.
- Supervised and unsupervised classification techniques based on extended ICA algorithms are proposed for brain tissue analysis.
- A new hybrid feature selection method using ICA and multisignal wavelet analysis is proposed for improved classification and segmentation using SVM.
- Limitations of ICA in brain MRI analysis, especially the poor performance in retaining local features, are resolved through spectral angle mapping and multisignal wavelet analysis.
- A multispectral analysis system is developed to evaluate the effectiveness of proposed methods and existing techniques in quantification of normal tissue and lesions. The results obtained from the experiments, using both unsupervised and supervised methods demonstrate that, with the proposed feature extraction approaches, accuracy of brain tissue analysis increases.

1.4.2 Clinical contributions

- Efficiency of the proposed methods in simultaneous analysis of small abnormalities, and their effect on other tissues and total brain volumes can help doctors in disease progress evaluation and treatment.
- Simultaneous analysis of multiple MR sequences reduces the time and effort required for clinical diagnosis.
- Computer aided diagnosis using new algorithms avoids inter/intra operator variability in pathological analysis, and it helps neuro-

radiologists and analysts to improve the accuracy and reliability of the final medical reports.

1.5 Thesis overview

In this chapter we outlined the problems associated with the brain MRI analysis, and highlighted the techniques that are going to be introduced in this thesis. The remaining chapters are organized as follows:

Chapter 2: Literature Survey presents a survey of the work related to the brain MRI analysis. Previously developed systems for segmentation and classification are reviewed. Results, merits, and limitations of these systems are also discussed. Multispectral MRI analysis techniques are emphasized. Works on feature extraction and classification using ICA and its extensions are also illustrated in detail.

Chapter 3: ICA and Brain MRI firstly, introduces the concept, theory and algorithms involved in ICA. Secondly, it highlights advantages and drawbacks of ICA in multispectral brain MRI analysis. Chapter 3 concludes with some case studies based on ICA extensions.

Chapter 4: Proposed ICA Extensions for brain MRI discusses three new ICA extensions to improve the classification performance of brain tissues from multispectral MRI; a new Spectral Clustering ICA (SC-ICA), a new Multisignal Wavelet ICA (MW-ICA) and modified Multiresolution ICA (MICA). A detailed explanation of each, including relevant theory, algorithm and examples are provided in this chapter.

Chapter 5: Brain MRI Analysis with Proposed ICA Extensions evaluates proposed ICA extensions and conventional ICA, with supervised and unsupervised classification approaches through quantitative and qualitative analysis. First, it describes the input datasets, data preparation using image registration and feature collection, and bias selection through different

validation techniques. After that, a detailed quantitative and qualitative analysis of results from supervised classification using SVM, and unsupervised classification using FCM is presented.

Chapter 6: A hybrid approach to brain tissue classification introduces the newly proposed hybrid feature selection method for improved brain tissue classification using SVM. To ensure the positive impact of the new method in identifying abnormalities and their effect on other tissues, quantitative and visual results from a detailed performance analysis is also discussed in it.

Chapter 7: Comparative Analysis and Discussion presents an overall summary of the current work, and comparative analysis of the proposed methods with conventional classification methods. Chapter 7 ends with an elaborate discussion on merits and demerits of spectral based approach and wavelet based approaches.

Chapter 8: Conclusions and Future Works recapitulate the thesis, and points out some possible extensions of the current work.

Chapter -2
LITERATURE SURVEY

2.1 Introduction

Brain tissue classification and abnormality detection has great significance in treatment planning and disease progress evaluation. Computer aided systems can help clinical experts to achieve the fast and accurate results in this regard. Even though several automated analysis techniques claims its superiority in analysis of a particular disease using single or multispectral images, there is no clear evidence for the suitability of these methods for various applications. A detailed review of MRI analysis methods till 1995 is available from Clarke *et al.* (1995) [7]. Relative merits of single versus multispectral segmentations, feature extraction, supervised and unsupervised segmentation methods as well as tissue classification methods were illustrated. Image pre-processing approaches, multimodality registration methods, validation procedures, examples for MRI segmentation and problems associated with inter- and intra-observer variations were also presented in detail.

Another literature survey on CAD systems for brain diseases in MRI by Arimura *et. al.* (2009) [34] covered a critical review on methods used in disease based analysis of MR images from 1994 to 2009, giving more emphasis on the works in the duration 2004-2009. The review started with fundamental techniques used for CAD systems, including image processing and pattern recognition. Pre-processing, feature extraction, classification, segmentation and evaluation were explained in detail in this paper. Several

application examples of CAD approaches were introduced, including detection of abnormalities. They concluded the survey with several future directions for improved computer aided diagnosis.

In this chapter, only the relevant works related to our thesis work such as transform based feature extraction, single and multispectral MRI image classification and segmentation are included. ICA and its extensions proposed in the literature, which are relevant to this work, are also discussed.

2.2 Common software tools in brain MRI analysis

Brain MRI analysis has been emerged as a hot research topic in the last two decades, with emphasis on several automatic and semi automatic methods closing the gap between technical and medical disciplines. But most of the implementations were performed for research purpose only. The most widely used software packages in the neuroimaging community are,

- SPM (Statistical Parametric Mapping) [35, 36, 37], written by the Wellcome Department of Imaging Neuroscience at University College London, UK). It is an approach, based on probabilistic atlases combined with fuzzy clustering, for brain segmentation and tissue extraction. Tissue classification requires the images to be registered with tissue probability maps.
- FSL (FMRIB Software Library) [38, 39], from Analysis Group, FMRIB, Oxford, UK. It is a software library containing image analysis and statistical tools for functional, structural and diffusion MRI brain imaging data. It combines an approach based on intensity distributions analysis and deformable models [40] for segmentation.
- FreeSurfer, written by collaboration between the Massachusetts Institute of Technology and the Harvard University, USA [41, 42]; a combined approach based on watershed algorithms, deformable surfaces and brain

atlases [43]. The problem of MRI segmentation is focused on segmenting normal brains without apparent diseases into three tissues: GM, WM and CSF ([44]). It is time consuming compared to SPM and FSL; the processing of one scan consumes 12-24 hours in comparison with the 15-30 minutes required by FSL and SPM. Also, segmentation of brain images requires a statistical atlas to give prior information about the spatial position of different structures.

Tsang *et al*. (2008) [45] conducted a quantitative analysis and comparison of the segmentation algorithms in latest versions SPM5 and FSL for the whole brain, as well as for the subcortical region. SPM5 segmentation algorithm yielded the highest overlap with the ground truth GM tissue maps for both the subcortical area and the whole brain. Performance metrics observed by Shattuck *et al.* (2009) [46], in a validation study based on FSL and FreeSurfer with on-line segmentation validation engine reported that indexes values can still be improved for FSL and Freesurfer.

More importantly, the results proposed by these packages are based on normal CSF, GM and WM structures, and they are found to be not efficient in multispectral and pathological analysis. In this regard, problem-specific studies were found to be yielding promising results in brain tissue analysis.

2.3 Feature extraction

Feature extraction is the process of extracting abnormal and normal tissue specific features from the pre-processed images in such a way that inter-class variation is maximized and intra-class similarity is maximized. Methods for feature extraction can be divided into different categories based on the type of features, such as pixel intensity-based features (gray scale values of the pixels), calculated pixel intensity-based features (calculated MR parameters such as metrics relating to flow of contrast material, cerebral blood volume,

blood flow or blood oxygenation), edges and texture-based features, Transform based features etc. [7].

In this thesis work we focus on transform based feature extraction methods to extract a set of discriminative features which provide better classification of MRI images. In literature, various feature extraction methods have been proposed such as PCA [31], ICA [47], Fourier Transform [48] and Wavelet Transform [49-51].

Fourier transform is useful for extracting frequency contents of a signal, however it cannot be used for analyzing accurately both time and frequency contents simultaneously. In order to overcome this, wavelet analysis is proposed which captures both low-frequency and high-frequency information accurately. For the classification of Alzheimer's disease, Chaplot *et al*. (2006) [49] used Daubechies-4 (db4) wavelet of level 2 to extract the features from MRI.

Dahshan *et al.* (2010) [51] pointed out that the features extracted using db4 Wavelet were too large, and may not be suitable for the classification. They used Haar wavelet of level 3 for feature extraction and further reduced features using PCA before classification. Though PCA reduce the dimension of feature vector, it is a computationally expensive global transform. Multiresolution analysis using wavelet transforms and its new extension, curvelet transform [52], have proved its potential in medical image segmentation with accurate clinical results [32]. Curvelets were found to give superior performance in local geometric feature analysis, but wavelets still keep its position in 1-D signal analysis [32, 52].

In a recent study, quantitative analysis of brain tissue classification using ICA [29] confirmed the efficiency of ICA in clinical MRI applications. However, these global transforms often fails to keep the local characteristics

in classification results. Han and Li (2011) [53] used a multiresolution approach to retain the critical local information in micro array classification, while performing a global transform using ICA. A growing interest in the application of multidimensional wavelet analysis to MR data [54] motivated us to use multisignal wavelet decomposition [53, 55] with ICA to extract the best spectral signatures corresponding to each brain tissue, which will be discussed in detail in the following chapters.

2.4 Survey of supervised classification methods in brain MRI analysis

Classification is the step that deals with the labelling of the regions. In supervised MRI analysis, pixels are associated with anatomical tissues by an expert in the training phase. Classification models obtained after training are used to classify the new samples. Many learning methods have been used to analyze intensity based information from images, including ANN [56], PCA [57], and SVM [58].

2.4.1 Artificial Neural Networks (ANN) based methods

The application of Kohonen neural networks for image classification was explored by Messen *et al.* (2006) [59]. Some modifications of the conventional Kohonen neural network were also implemented in this work which proved to be much superior to the conventional neural networks. The different grades of abnormal images were categorized using ANN by Yamashita *et al.* (2008) [60]. This report suggested a practical method for selection of database. The training of ANN is dependent on input data and hence a wide variety of pattern is desirable for high accuracy. This report also highlighted the difficulty in collecting a large dataset of different uncommon patterns. Though the report records high classification accuracy, the size of the dataset is significantly small.

El-Dahshan *et al*. (2009) [61] proposed the hybrid technique consisting of three stages; feature extraction using Discrete Wavelet Transformation (DWT), dimensionality reduction by PCA and classification by Feed Forward Back-Propagation Artificial Neural Network (FP-ANN) and k-Nearest Neighbor (k-NN). Comparison with Self Organizing Map (SOM) and SVM revealed that ANN method gained the worst sensitivity and specificity rate. The index values were based on single image analysis with small number of datasets.

Jensen and Schmainda (2009) [62] explored different neural networks to detect brain tumor invasion from multi-parametric MRI (structural, diffusion and perfusion images). An automatic brain tumor detection method using Gabor wavelets was proposed by AmirEhsan Lashkari (2010) [63]. The neural network had been trained using back propagation algorithm and training process was continued until the Mean Square Error (MSE) became constant with about accuracy of 98.15%.

The system designed for brain cancer detection and classification by Joshi *et al.* (2010) [64] used conceptually simple classification method using the Neuro Fuzzy logic. Texture features are used in the training of the ANN. This system provides precision detection and classification of astrocytoma type of cancer. Zhang *et al.* (2011) [65] presented a neural network (NN) based method to classify a given MR brain image as normal or abnormal. DWT, PCA and Back Propagation Neural Network (BPNN) were used for classification. The classification accuracies on both training and test images were 100%, and the computation time per image was only 0.0451 seconds. It was a preliminary study with singe image analysis using T1WI.

2.4.2 Classification based on Support Vector Machines (SVM)

Majority of the recent works in Brain MRI classification were found to be based on SVM. Relevant methods using SVM are summarized in Table 2.1.

Table 2.1 Methods using SVM classification for MR image analysis

ID	Methods and description
Li et. *al*. (2006) [66]	Various levels of MR glioma images were classified using SVM. Better than fuzzy rule based systems, but the accuracy reported was low. It deals only with glioma images.
Chaplot *et. al.* (2006) [49]	A hybrid approach using wavelets and SVM for classifying the abnormal and normal images. Better than the neural networks in terms of performance measures. Small size dataset used for implementation. The classification accuracy results may reduce when the size of the dataset is increased
Luts *et al.* (2007) [67]	Least Square SVM (LS-SVM) used for brain tumor recognition. Both bi-level classification and multiclass classification were performed. Suggested the necessity for multiclass classification techniques than bi-level classification techniques.
Selvaraj *et al.* (2007) [68]	Least Square SVM (LS-SVM) for brain tumor recognition performed. An extensive comparative analysis results with SVM, neural classifier and the statistical classifiers suggested the advantages of SVM in terms of classification accuracy. Only bi-level classification is performed.
Chandra *et al*. (2009) [69]	RBF kernel based SVM for brain tumor detection used. Results compared with AdaBoost machine learning algorithm. Superior nature of SVM over the other classifiers illustrated.
Hamilton - Wright *et al.* (2009) [70]	Image classification based on fuzzy approach using the pattern discovery algorithm. Better results observed when compared with the other classifiers
Zacharaki, *et al.(*2009) [71]	Pattern classification methods for distinguishing different types of brain tumors, such as primary gliomas from metastases, and also for grading of gliomas. Feature subset selection was performed using SVM with recursive feature elimination. High performance classification provided. Failed to incorporate features describing the deformation of the healthy structure due to tumor growth.
Ruan *et al.* (2011) [72]	Claimed that the feature selection using kernel class separability could slightly improve the results.

Zhang and Wu (2012) [73]	Wavelet transform for feature extraction, followed by PCA to reduce the dimensionality. A kernel support vector machine (KSVM) used for classification. K-fold stratified cross validation was used to enhance generalization of KSVM. The highest classification accuracy, 99.38% observed, but with a small size dataset.
Mubashir Ahmad *et al.* (2012)[74]	Db4 wavelet transform, PCA and SVM with Linear Kernel and Radial Basis Kernel were used in the validation. Experimental results showed high classification accuracy of 98.7% with Radial Basis Kernel.

Compared with other methods such as artificial neural network, decision tree, and Bayesian network, SVMs have significant advantages of high accuracy, elegant mathematical tractability, and direct geometric interpretation. Besides, it does not need a large number of training samples to avoid overfitting. The dominant feature which makes SVM very attractive is that classes which are nonlinearly separable in the original space can be linearly separated in the higher dimensional feature space.

In addition to these conventional methods, other approaches yielding high performance in brain MR classification are PNN and Bagging. The modified PNN for tumor image classification is used by Georgiadis *et al*. (2008) [75]. Abnormal images such as metastases, glioma and meningioma were differentiated using the least square feature transformation based PNN. A comparative analysis is also performed with SVM. This work inferred that the transform based PNN is superior to the SVM in terms of classification accuracy. Ibrahiem and Ramakrishnan (2008) [76] stated that a time efficient neural network such as PNN can be used for pattern classification problems. Emphasis was given for convergence time than the classification accuracy. The results concluded that the PNN is superior over conventional neural networks in terms of training time period.

Bagging, which stands for bootstrap aggregation[77] is another way of manipulating training data for ensemble classification schemes. In each

iteration, the training subset is bootstrapped (resampled with replacement), to generate a different training subset. The logic behind bagging is that unstable classifiers, such as neural networks and decision trees, whose behaviour could be significantly changed by small fluctuations in the training subset, are more likely to be stabilized after being trained with different input data [78]. Although the SVM classifier has been shown to provide a good generalization performance, the classification result of the SVM is often far from the theoretically expected level, because SVM implementations usually employ approximation techniques [79].

2.5 Survey of brain MRI image segmentation approaches

MR image segmentation refers to the decomposition of an image into regions representing the different tissue types. In the case of supervised MRI analysis, classification models are used for automatic tissue segmentation. Unsupervised classification generates segmented tissues as the first results, upon which the classification is performed. Actually, it is a pixel based classification, since the individual pixels are clustered unlike the classification techniques which categorize the whole image. Several research works are reported in the area of MRI image segmentation.

Clustering was introduced into the brain tumor segmentation community by Schad *et al* (1993) [80], who analyzed texture patterns of different tissues. Phillips *et al* (1995) [81] employed FCM, and Vaidyanathan (1995) [82] compared this to k-NN clustering for the tumor volume determination during therapy on multi-spectral 2D image slices. Clark *et al* (1998) [83] from the same group, further developed this approach to incorporate knowledge-based techniques. A complete survey of various segmentation algorithms is presented by Dzung and Jerry (2000) [84]. The merits and the demerits of various techniques were analyzed in detail. Appropriate techniques for

different applications were also suggested in their review paper. An extensive survey on tissue segmentation algorithms for MR brain images was conducted by Liew and Yan (2006) [85]. A technique to minimize the intensity nonuniformity artifact was also proposed in that paper. Several suggestions on pixel based approaches, region based approaches, model based approaches were provided in their work. Some relevant works we observed based on pattern recognition techniques are summarized in Table 2.2.

Table 2.2 General pattern recognition methods in Brain MRI segmentation

ID	Methods and description
Fletcher-Heath *et al.* (2001) [86]	Fuzzy clustering with knowledge based techniques for the brain tumour segmentation from multi-sequence MRI.
Moon *et al.*, (2002) [87]	A model based tumor segmentation technique using a modified Expectation Maximization (EM) algorithm. Lack of quantitative analysis on the extracted tumor region observed.
Zhu and Tian (2003) [88]	Level set method, which involves the method of boundary detection with the seed point. Watershed algorithm used to capture the weak edges. Random selection of seed point leads to inappropriate results and also consumes large convergence time period.
Li and Chi, (2005) [89]	A modified version of SOM with Markov random field model. Extra spatial constraints added in this algorithm for weight adjustments. This method was highly prone to noise and applicable for only noise-free images.
Deorah *et al.*, (2006) [90]	A complete analysis of various types of brain tumors and the effect of MR image segmentation techniques on the treatment.
Zaidi *et al.* (2006) [91]	The merits and demerits of various statistical segmentation techniques discussed. Analyzed the performance measures of histogram based method, EM technique and the SPM package in detail.
Martin-Landrove and Villalta (2006), [92]	Back Propagation Neural Network (BPNN) based tumor segmentation. A comparative analysis was done with the Inverse Laplace Transform based technique. Concluded that BPNN is superior in terms of processing time and accuracy over the conventional algorithm.
Khotanlou *et al.* (2007) [93]	An enhanced version of symmetry analysis which also incorporates the deformable models. The segmentation efficiencies was very low, and a

	failure in case of symmetrical tumor across the mid-sagittal plane.
Unnikrishnan *et al.* (2007) [94]	Conventional algorithms such as EM algorithm, mean shift filtering algorithm, etc. were experimented. Comparative analysis between various performance measures was not reported.
Ray et al. (2008)[95]	Symmetry based brain tumor extraction. It can detect only the densely packed tumor tissues.
Lee *et al.* (2008) [96]	Application of pseudo-conditional random fields for brain tumor segmentation demonstrated. Claimed to be highly accurate and much faster than other conventional techniques. Patient - specific training was used.
Rasoul Khayati *et al.* (2008) [97]	Bayesian classifier, utilizes the adaptive mixtures method (AMM) and Markov random field (MRF) model to obtain and upgrade the class conditional probability density function (CCPDF) and the a priori probability of each class. High performance segmentation observed.
Corso *et al.* (2008) [98]	Bayesian model based tissue segmentation technique for tumor detection. Computationally efficient besides yielding improved results over conventional techniques.
Anbeek *et al.* (2008) [99]	K-Nearest Neighbor technique used. An extensive comparative analysis performed with other techniques. The dependency on threshold values for accurate output was the drawback.
Ma *et al.* (2009) [100]	A survey on various medical image segmentation algorithms. The drawbacks of several techniques were clearly illustrated, and suggested suitable techniques for tumor segmentation.
de Boer *et al.* (2010) [101]	Evaluated the accuracy and reproducibility of four previously proposed automatic brain tissue segmentation methods: FAST, SPM5, an automatically trained k-NN classifier, and a conventional k-NN classifier based on a prior training set. The conventional k-NN classifier method performed best in the accuracy experiment and worst in the reproducibility experiment. FAST showed the best reproducibility, but its accuracy was relatively low for CSF and GM. Overall, the accuracy and reproducibility were good, and there were only small differences between the methods.
Kanaly *et al.* (2011)[102]	A simpler approach by thresholding the voxels of the difference image of pre- and post-contrast T1-weighted MRI after intensity normalization.

Data mining algorithms like KNN and FCM are often used in brain tissue segmentation. In addition, the iterated conditional modes or some neural network methods like SVM can be applied for this problem. However, it is difficult to compare the reported accuracies of these segmentation methods. Different evaluation measures were used and some of these measures depend on the tissue volumes of the subject, and the evaluation was influenced by the manual segmentation protocol.

Recent studies demonstrated FCM as a good technique for unsupervised classification. They are robust to ambiguity and efficient in retaining much more information than hard segmentation methods. We reviewed a few relevant works based on fuzzy approaches as a part of this thesis work, as given in Table 2.3

Llado *et al*. (2012) [24] have done a critical review of segmentation of multiple sclerosis lesions using common approaches such as *A*NN, kNN, AdaBoost, Bayesian classifiers, and FCM. The main features of the segmentation algorithms were analysed, and the most recent important techniques were classified into different strategies according to their main principle, pointing out their strengths and weaknesses. They observed that the automated segmentation of different MS lesion types in MRI is a challenging task due to heterogeneous intensity values among the different MR images (enhancing lesions, black holes and hyper intense lesions). They observed that there is not yet a specific automated lesion segmentation approach robust enough to emerge as a standard for clinical practice. The main reasons are unsatisfactory results they produce, the high computational demand required, or their insufficient generalization capability.

Table 2.3 Fuzzy approaches in brain MRI segmentation

ID	Methods and description
Cheng *et al.,* (1998) [103]	A modified Fuzzy C-Means (FCM) algorithm for brain image analysis. It provided significant time saving when compared with the conventional FCM algorithm. Lack of quantitative analysis on segmentation efficiency is the drawback.
Dzung &Jerry (1999) [84]	More accurate FCM algorithm. Applicable for MR images with intensity inhomogeneties. Execution time is directly proportional to the amount of inhomogeneties present in the images. Slower convergence rate observed.
Boudraa *et al.,* (2000) [104]	FCM preceded by a local image contrast enhancement procedure. Reduced false positives and false negatives observed. Reproducibility is to be evaluated with large size database.
Dave & Sen (2002) [105]	A robust fuzzy clustering algorithm. Eliminated the dependency of FCM algorithm on similarity measures such as distance measures. Highly generalized in terms of the parameters used in the algorithm.
John & Hutcheson, *et al.* (2002) [106]	Time efficient FCM for real time applications. Solved the high memory requirement problem of conventional FCM.
Khalighi *et al.,* (2002) [107]	A modified FCM algorithm based on spatial model. Qualitative results highlighted the improvement in the segmentation efficiency of the spatial model based FCM over the conventional FCM. The lack of volumetric analysis of the segmented tissues is the major drawback.
Eschrich *et al.,* (2003) [108]	A fast and accurate FCM algorithm. A data reduction method based on quantization is involved in this approach for high speed clustering.
Zhang & Chen, et al., (2004) [109]	Kernelized fuzzy C-means (KFCM) algorithm in which Euclidean distance in the FCM is replaced by a kernel-induced distance. Spatial penalty is added to objective function in KFCM to compensate for the intensity inhomogeneities. Better performance when noise and other artifacts are present.
Kannan (2005) [110]	A FCM algorithm to achieve high accuracy for MR brain images. Only qualitative analysis is reported in the paper which is not sufficient for judging the effectiveness of the system.
Murugavalli & Rajamani (2006) [111]	An improved FCM algorithm based on block processing where each block is processed by a parallel processor. Though this approach is faster, the requirement for additional hardware is the major drawback.
Dou *et al.* (2007)	Several fuzzy models were created and the fuzzy features were

[112]	extracted from these models. Experimental results suggested that usage of fused fuzzy features for improving the accuracy of the techniques. An extensive analysis in terms of the measures is required.
Juang *et al.* (2007) [113]	A combinational approach of SOM, SVM and fuzzy. An extensive analysis performed with the other segmentation techniques to show the superior nature of the proposed approach.
Kannan (2008) [114]	Modified FCM algorithm with method of initial selection of membership values highlighted. Significant experiments not conducted which is evident from the lack of quantitative analysis.
Zhou *et al.* (2008) [115]	Two-dimensional FCM clustering algorithm. The simulation results showed the segmentation accuracy, 98%. Algorithm has strong anti-noise capability, high clustering accuracy and good segment effect.
Ramakrishnan *et al.* (2010) [116]	Fuzzy multi wavelet packet transformation based brain MR image classification method. Experiments showed that the fuzzy-based criterion achieves higher recognition rate with relatively smaller sub bands than signal energy-based criterions. Higher recognition under noisy environment with lesser number of sub bands achieved.
Yang & Fei (2011) [117]	An automatic, multiscale and multiblock fuzzy C-means (MsbFCM) classification method with MR intensity correction. Artifacts like noise and intensity inhomogeneities addressed. Accurate and robust for various MR images
Kang *et al.* (2011) [118]	Intelligent generalized tissue classification system which combines both the Fuzzy C-means algorithm and the qualitative medical knowledge on geometric properties of different tissues
Iraky khalifa *et al.* (2012) [119]	Wavelet Fuzzy C- means (WFCM) algorithm. Feature extraction using multilevel 2D wavelet decomposition with db4 was used to do FCM clustering. A detailed analysis is not provided to justify the improvement in real cases.

2.6 Survey of methods in multispectral brain MRI analysis

For the last two decades, simultaneous extraction and analysis of relevant and complementary information using multispectral analysis of different MRI sequences has been a widely discussed research topic [4, 7, 22]. MRI sequences like T1WI, T2WI, PDI, FLAIR images etc. provides a huge

repository of unique information on different tissues [4, 7]. For example, T1WI shows considerable contrast between GM and WM. T2WI can give details of CSF and abnormalities; sometimes it cannot distinguish between CSF and abnormalities like edema and lesions. Then FLAIR images, which suppress CSF effects to give more details on hyper-intense lesions, can be utilized to separate the normal and abnormal tissues. Simultaneous analysis of each sequence to collect the prominent pathological information is a tedious job for radiology experts. Computer aided diagnosis using multispectral approach is helpful in this context to save time, and to improve the accuracy and consistency of the clinical results [34]. Conventional algorithms in the literature for single image analysis are not efficient and robust to yield good results with expected clinical accuracy for multispectral data. Feature extraction and optimal feature selection can reduce the computational overhead, and improve the overall performance. As an application of pattern recognition, multi-spectral MR image segmentation methods are subcategorized into unsupervised and supervised approaches.

2.6.1 Unsupervised approaches

For multi-spectral medical image segmentation, unsupervised methods have the advantage of avoiding the training process, which includes human operation variability associated with manual selection of training samples [7]. From the literature, it was seen that the most widely used FCM algorithm [120, 121] has been consistently improved in the past decades to achieve better performance in MRI segmentation results. Its simplest form, *k*-means [121] is also found to be useful in preliminary studies. Expectation Maximization (EM) has also been used very frequently in multi-spectral image segmentation [122, 123]. It applies the same clustering principles as that of FCM and *k*-means algorithms, with an assumption that the input image data follow a Gaussian Mixture Model (GMM).

Normal structure segmentation from multi-spectral MR images involves only the known tissues like CSF, GM, WM etc., where unsupervised segmentation methods were found to be generating satisfactory results [7, 19]. However, cluster is not a well defined concept [124] to group unknown classes like different abnormalities, which makes it undesirable for abnormal MRI analysis. A patient-specific training is necessary in the pathological analysis to categorize the different lesions and tumors of varying sizes.

2.6.2 Supervised multispectral analysis

Supervised methods can automatically identify the relevant patterns representing different abnormalities in the data, based on prior knowledge collected from the history of different abnormal cases [7]. The most frequently used supervised pattern recognition techniques for medical image segmentation included the Maximum Likelihood (ML) [22], *k*-NN algorithms [125] and ANN [7] in the past. Later, studies showed that ANN [26], PNN [75], SVM [30], and data conditioning approaches were capable of yielding high performance classification in multispectral analysis.

The ML method [126] is a successful approach when the input data distributions for the different classes are well known. But, for multi-spectral images, it was found to be impractical [7]. The *k*-NN algorithm is a nonparametric approach, which does not rely on predefined input data distributions, but it depends on the actual distribution in the training samples [127].

The pixel based SVM method in multispectral MRI analysis, which has seen much popularity in last decade [128], showed high performance with fair results. Later, Verma *et al* (2008) [129] used a high number of MRI modalities (diffusion tensor imaging (DTI) channels in addition to the conventional ones) to create voxel-wise intensity-based feature vectors,

which they classified with Support Vector Machines (SVMs). They were able to not only segment the healthy tissues, but also segment sub-compartments of healthy and tumor regions.

2.6.3 Recent developments

In recent years several studies have demonstrated the efficiency of multispectral analysis in brain tissue classification and abnormality analysis. Since Clarke *et al.,* (1995) [7] and Y. Kvinnsland *et al.* (2009) [4] covered a comprehensive review on almost all multispectral methods in MRI, following discussion includes only some most relevant and recent studies in this field, and a summary is given in Table 2.4.

He *et al.* (2008) [6] proposed a Generalized fuzzy clustering approach based on FCM (GFCM) for segmentation of multispectral MR images. A bias field correction and contextual constraints over spatial intensity distribution were included in it to account for the non-spherical cluster's shape in the feature space. The bias field was modeled as a linear combination of smooth polynomial basis functions for fast computation in the clustering iterations. Since the feature space was not isotropic, distance measure adopted in Gustafson-Kessel (G-K) algorithm [130] was used instead of the Euclidean distance, to account for the nonspherical shape of the clusters in the feature space. They observed that conventional FCM-based methods did not correct the intensity inhomogeneity and did not exploit the contextual information. Qualitative and quantitative evaluation of normal brains indicated that GFCM outperforms the conventional FCM method. The performance of GFCM is to be explored with additional dimensions (such as FLAIR images) in the feature space for a complete study of the separability of more classes such as lesions.

Kroon *et al.* (2008) [131] presented a local feature vector based method for automated Multiple Sclerosis (MS) lesion segmentation of multi-spectral MRI data. Local feature vector method in the study included neighborhood voxel intensities, histogram and MS probability atlas information. PCA with log-likelihood ratio was used to classify each voxel. Bias field corrections based on genetic algorithm, edge preserving filtering and atlas based correction were introduced to intensity inhomogenities in MR images. The similarity scores between proposed model and expert classifications were found to be less. Best classification results were observed with normal raw data because all the described bias correction methods caused artifacts at region edges.

Lao *et al.* (2008) [132] presented a computer-assisted WML segmentation method, based on local features extracted from MRI sequences, T1-weighted, T2-weighted, proton density-weighted, and FLAIR MR scans. Although FLAIR provides the best contrast between periventricular WMLs and ventricles, PD helps to avoid potential overestimation of lesion load that has been observed with the FLAIR sequence The proposed approach started with a number of pre-processing steps; intrasubject co-registration, skull stripping, inhomogeneity correction, and intensity normalization. In the learning process of the method, optimal generalization ability of SVM was highly exploited. AdaBoost[77] approach helped to learn progressively from misclassified examples. Automated removal of false positives was also incorporated in the method. Cross-validation on a population of 35 patients confirmed the robustness and accuracy of the proposed segmentation method, compared to the manually segmented results by two experienced neuro radiologists. A rigorous evaluation of relative value of each acquisition protocol was not performed in the study.

Lecoeur *et al.* (2009) [133] presented an optimized supervised segmentation method for multispectral MR analysis. An optimized spectral gradient based on a psycho-visual graph cut paradigm was created using multispectral MR analysis. Using Dice similarity coefficient as a cost function for an optimization algorithm, an optimized gradient was computed and it was utilized to segment MRIs with the same kind of modalities. Results showed that the optimized gradient matrices yielded significantly better segmentations and that the supervised learning of an optimized matrix is a good way to enhance the segmentation method. Working of this scheme on different combinations of sequences for improved pathological analysis was also demonstrated.

Ozer *et al.* (2010) [134] developed automated methods that combine the pharmacokinetic parameters derived from dynamic contrast enhanced (DCE) MRI with quantitative T2W MRI and diffusion weighted imaging (DWI) in contrast. Large margin classifiers were used for prostate cancer segmentation. New thresholding schemes were developed to tune SVM and their probabilistic counterparts, Relevance Vector Machines (RVMs), for an improved performance with respect to a selected criterion. Moreover, a thresholding method was applied to make a fully automatic unsupervised fuzzy Markov random field method. No significant difference between the SVM and RVM segmentation results was observed. The results of the automated algorithms indicated that multispectral MRI analysis improved prostate cancer segmentation performance compared to single image analysis. Multispectral MRI analysis provided better information about cancer and normal regions in the prostate, which were found to be very useful in the diagnosis and detection of prostate cancer. The analysis results confirmed the superiority and robustness of the supervised methods in terms of standard deviation when compared to unsupervised methods.

Zhang *et al.* (2011) [135] presented a framework of medical image analysis system for the brain tumor segmentation, and the brain tumor following-up over time using multispectral MRI images. Multispectral images have the advantage of providing complementary information to resolve some ambiguities in brain tumor analysis. However, they may also bring along a lot of redundant information, increasing the data processing time and segmentation errors. The SVM classification integrated with a selection of the features in a kernel space was proposed in the work to resolve this issue. The selection criteria were defined by the kernel class separability. A brain tumor evolution follow up based on the same SVM classification framework was also proposed. The system has been tested on real patient images with satisfying results. The quantitative evaluations by comparing with manually traced results and with other conventional approaches demonstrated the effectiveness of the proposed method. An elaborate validation of whole process in a much larger database with different patients over long therapeutic periods is required to establish the accuracy and robustness in disease progress evaluation.

Valdés Hernández *et al.* (2012) [8] proposed a novel unsupervised classification method for WML analysis, Multispectral Coloring Modulation and Variance Identification (MCMxxxVI). Two different structural Magnetic Resonance Imaging (MRI) sequences in red/green color space were first fused, and Minimum Variance Quantization (MVQ) [136] was applied as the clustering technique to segment the different tissue types. Then they investigated its performance compared with several commonly used supervised image classifiers in segmenting normal-appearing white matter, white matter lesions and cerebrospinal fluid in the brains of 20 older subjects. In the multispectral analysis, they segmented these three tissue classes from T1W, T2W, T2*W and fluid attenuation inversion recovery (FLAIR)-

weighted structural MRI data using MCMxxxVI and the four supervised multispectral classifiers available in the Analyze package [137], namely, Back-Propagated Neural Networks, Gaussian classifier, Nearest Neighbour and Parzen Windows. Bland–Altman analysis [138] and Jaccard index [46] were used in performance improvement evaluation. Results indicated that, in general, MCMxxxVI performed better than the supervised multispectral classifiers in identifying the three tissue classes. However, final manual editing was required to deliver clinically acceptable results. The two limitations of this study were lack of an explicit investigation on inter-observer variations and relatively small size of sample of subjects studied.

García-Lorenzo *et al.* (2013) [139] presented a systematic and latest comprehensive review of the literature to evaluate the state of the art in automated multiple sclerosis lesion segmentation. They stated that many methods provide limited solutions, where they deal with only one type of MR protocol or identify only one type of MS lesion, and they rarely address the complexity of the diseases. Nevertheless, many advances have been made over the years, and many methods have demonstrated promising results with MRI data from small groups of patients. According to their paper, the challenge remains to provide segmentation techniques that can work in all cases regardless of the type of MS, duration of the disease, or MRI protocol, and within a comprehensive, standardized validation framework. They described the complexity of the lesion segmentation problem and identified solutions that have been proposed in the literature. In addition to describing the techniques, they focused on the strengths and weaknesses of the validation methods used to characterize the published methods. They concluded with a few suggestions for improvement and good validation. Conclusions and suggestions from their review are repeated in the next section in a general point of view for MRI analysis.

Table 2.4 Summary of recent developments in multispectral MRI analysis

ID & Title	Input data	Methods	Results	Remarks
He *et al.* (2008) [6] Generalized Fuzzy Clustering for Segmentation of Multi-Spectral Magnetic Resonance Images	T1W, T2W, ,PD	A comprehensive GFCM technique	GFCM outperforms the conventional FCM method and its extensions in normal brain MRI analysis.	Study included only normal brains. More detailed study using additional dimensions (such as FLAIR images) required.
Kroon *et al.* (2008) [131] Multiple sclerosis detection in multispectral magnetic resonance images with principal components analysis	FLAIR, T1W, T2W, and raw data	PCA with log-likelihood ratio, Genetic algorithm, edge preserving filtering and atlas based correction used for bias field correction	The similarity scores between proposed model and expert classifications were found to be less. Best classification results observed with normal raw data.	Better local features and separate PCA reductions for MS and non-MS data required. A larger training set also required to improve the results.
Lao *et al. (2008)* [132] Computer-Assisted Segmentation of WML in 3D MR Images Using Support Vector Machine	T1W, T2W, PD, FLAIR	SVM	Post processing analysis reduces false positives by using anatomic knowledge and measures of distance from the training set. Did not rigorously evaluate the relative value of each acquisition protocol.	Divided patients by lesion load. Quite good result and did the evaluation of binary segmentations

Lecoeur *et al.* (2009) [133] Optimized supervised segmentation of MS lesions from multispectral MRIs	T1W, T2W, PD, FLAIR	An optimized spectral gradient based on a psycho-visual graph cut approach	Improved pathological analysis with overall lower computational load	Validation using larger datasets required to ensure the reliability and robustness
Ozer *et al.* (2010) [134] Supervised and Unsupervised Methods for Prostate Cancer Segmentation with Multispectral MRI	T2, DWI, and DCE-MRI	SVM, RVM, fully automatic unsupervised fuzzy Markov random field method	No significant difference between the SVM and RVM segmentation results Improved prostate cancer segmentation performance compared to manually segmented results.	Found to be very useful in the diagnosis and detection of prostate cancer. Recommends supervised methods for brain MRI analysis.
Zhang *et al.* (2011) [135] Kernel feature selection to fuse multi-spectral MRI images for brain tumour segmentation	T1W, T2W, PD, FLAIR	Support Vector Machine (SVM) integrated with kernel feature space, adaptive training process	Compared with experiments without feature selection and adaptive training, the proposed method improved the results significantly and reduced the time consumption.	Self-updated system over time provides latest results. An elaborate validation required for larger database
Valdés Hernández *et al. (*2012) [8] Automatic segmentation of brain white matter and white matter lesions in normal aging: comparison of five multispectral	T1W, T2W, T2*W, FLAIR	MCMxxxVI, Back-Propagated Neural Networks, Gaussian classifier, Nearest Neighbour, Parzen Windows.	MCMxxxVI delivered better results than the other methods investigated	Lack of an explicit investigation on inter-observer variations, relatively small size sample of subject studied. Methodological improvement required for full automation

techniques.				
García-Lorenzo *et al.* (2013) [139] Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging		Systematic review of the literature to evaluate the state of the art in automated multiple sclerosis lesion segmentation	Described the complexity of the lesion segmentation problem and identified solutions that have been proposed in the literature	Latest and a good review in the research of brain MRI segmentation

2.7 Challenges in MRI analysis

Challenges in MR image analysis can be divided into two categories [139]; First, issues related to the proposed methods and implementations, and second, problems related to characteristics of data.

2.7.1 Methodological challenges

A robust, accurate, fully-automated lesion segmentation method suitable for use in clinical trials is still not available: Available literature on the topic of brain MRI analysis and lesion segmentation, and some important advances have been described in the previous sections and survey papers. But it is clear that the problem is far from solved, and still more work is required in this area.

Importance of Multimodal/Multispectral information: Although some lesions are obvious in only one sequence (for example in FLAIR), the lesion should be confirmed using other sequences (T2W or PD) to avoid false positives.

T1W images also can give very good tissue contrast, which is found to be very useful to provide an improved analysis.

Spatial information is necessary: Both spatial and intensity information is important for a good segmentation. In some MR sequences more than one tissue have similar intensities; then noise will reduce the performance of the algorithm. Spatial information in the local neighborhood level and the anatomical level is useful in this regard resolve this issue.

Unsupervised vs. supervised approaches: Unsupervised methods have the advantage of not requiring a ground truth and prior knowledge about the brain tissue classes. But studies in the literature concluded that if the supervised methods are well trained, more robust results can be obtained. The majority of unsupervised methods are based on global clustering techniques such as FCM or a Gaussian mixture model. The intensity of the tissues and lesions vary across the image. So, supervised and unsupervised methods should consider these local variations in order to improve the sensibility of the methods in some parts of the brain.

Availability of methods: The majority of methods are not publicly available, which makes a comparative study and improvement very difficult.

2.7.2 Data specific challenges

Partial volume effects: Limited resolution of MR images lead to fuzziness in the border of the lesions, and it causes the partial volume effects. Methods considering these effects can improve the reproducibility of WML segmentation.

Multicenter datasets: Images from different scanners have different contrasts or intensities, even though same protocols are employed. Methods should be designed specifically to deal with this variation also, without affecting the significant information relevant for clinical studies.

Diffuse disease: Automatic methods have concentrated on the focal lesions caused by the disease, but in some cases it is impossible to find the border between lesions and the neighboring diffuse dirty WM. No method has attempted to resolve it; still it remains as an open issue.

2.8 Survey of ICA based approaches in biomedical applications

2.8.1 Conventional ICA

Blind source separation (BSS) is the approach, in which original source signals are estimated using only the information of the mixed signals observed in each input channel, where the independence among the source signals is mainly used for the separation. The early contributory works on BSS were performed by Cardoso (1989) [140] and Jutten and Herault (1991) [141], where higher-order statistics of the signals were used for estimation of the independence. Common (1994) [142] clearly defined the term *independent component analysis* (ICA), with the help of a new algorithm that measures independence among the source signals. Later, Bell and Sejnowski (1995) [143] followed this work to design ICA extension to the infomax (or the maximum-entropy) algorithm for BSS, which is based on a minimization of mutual information of the signals. Some relevant works related to brain tissue classification from multispectral data is explained in this section. More details on the theory, algorithms, applications and relevance of ICA in this thesis work are explained in Chapter 3.

ICA has shown its great ability in enhancing the image contrasts of major brain tissues in structural MRI and in analysis of functional MRI [144]. It is a potential and widely accepted method in fMRI analysis, but relatively less number of works was found in structural MRI analysis. Nakai *et al.* (2004) [33] proposed ICA as an unmixing method to separate different brain tissue

structures like GM, WM, and CSF into a set of statistically independent components. ICA based on the learning rule of Bell and Sejnowski after pre-whitening operations was used for evaluation. The involvement of gray or white matter in brain tumor cases and the demyelination in the case of multiple sclerosis were enhanced and visualized in independent component images. They suggested that with the proper choice of contrast for the original images, ICA is useful for preprocessing transformation before clustering and segmentation of human brain. This work was later widely accepted as the base for ICA based MRI analysis.

Certain limitations observed in the implementation of their work are given below:

- They assumed that the number of input MR bands, *L*, is greater than or equal to the number of brain tissue sources *p*. In contrast to fMRI, L< p (over-complete) for structural MRI analysis.
- Another issue is, random initial projection vectors by ICA, which results in different final sets of projection vectors and ICs for the same users in different runs or different users at the same time. This serious inconsistency undermines repeatability of ICA and makes it unstable.
- Additionally, due to the use of random initial projection vectors, the order of ICA-generated ICs is completely random and does not necessarily indicate the significance or importance of an IC. So, image evaluation must wait until all ICs are generated.

Some relevant works related to application of ICA in biomedical signals is well explained in James and Hesse (2005) [145]. Later, Chen and Sugiki (2006) [146] proposed a method to extract three independent components from MR datasets by ICA. The ICs obtained were used for improved segmentation of phantom MR images, T1W, T2W and PD images. Tateyama

et al. (2007) [147] proposed a new method for classification of brain matters in the MR datasets using Kernel Independent Component Analysis (KICA) [148] to address the non-linearity issues in input MR data. The experimental results showed that performance of the proposed method significantly improved compared to the conventional methods. But they did not perform a complete validation and analysis using a large database. Detailed explanation of this method with examples is available in Chapter 3.

Wenlu Yang *et al.* (2010) [149] proposed a method of automatic classification of magnetic resonance images based on ICA, which was composed of three steps. First, all MRI scans were aligned and normalized by SPM. Then FastICA was applied to the pre-processed images for extracting specific neuroimaging components as potential classifying feature. Finally, the separated ICs were fed into a classifying machine that discriminates among Alzheimer's patients, and mild cognitive impairment. The experimental results showed that this method can successfully differentiate subjects with Alzheimer's disease and mild cognitive impairment from normal controls. But vagueness in the meaning and usage of disease related features restricted them to generalize the method.

2.8.2 Solutions to Over-Complete ICA

Ouyang *et al.* (2008) [5] proposed an approach to solve Over Complete-ICA (OC-ICA) issue, and issue of random initial projection vectors. They implemented OC-ICA in conjunction with spatial domain-based classifications, to improve the classification performance. Their work was considered as the first attempt to investigate the utility of the OC-ICA in MR multispectral image analysis. Three ICA-generated independent components (ICs) were stacked one atop another to form a new image cube, which is spectrally and statistically independent in ICs. Proposed technique was observed as the best compromise between using supervised classification

alone such as Fisher's Linear Discriminant Analysis (FLDA) and SVM, and unsupervised technique alone such as ICA. Classifications using FLDA and SVM solved the inconsistency in ICs, and results indicated significant improvement in classification performance also.

The same team proposed another efficient method in [150], to address the same ICA issues. In order to address the over-completeness issue, they proposed a Band-Expansion Process (BEP), in which an additional new set of images were generated from the original MR images via nonlinear functions. These newly generated images were then combined with the original MR images to provide sufficient MR images for ICA analysis. In order to resolve the second issue, a Prioritized ICA (PICA) was designed to rank the ICA-generated independent components (ICs) so that MR brain tissue substances can be unmixed and separated by different ICs in a prioritized order. Finally, BEP and PICA were combined to further develop a new ICA-based approach, PICA-BEP, to perform MR image analysis using SVM. Prioritized ICA (PICA) was originated from dimensionality reduction concepts, eigenvectors of the data matrix and the higher-order statistics referred to as skewness and kurtosis were used in it. Even though prioritized ICA-BEP improved the traditional ICA, it was observed that generation of more bands may result in over separation in the sense that a brain substance is forced to be split and separated in more ICs. For three-band MRI analysis, it was recommended that using cross correlation may be good enough. More explanations on prioritized ICA-BEP with examples are included in Chapter3.

In another work, Chai *et al.* (2009) [151] compared two new approaches, ICA+SVM and BEP+ICA+SVM, to a standard classification technique, SVM. The results from synthetic data analysis for GM and WM classification showed improved results for BEP+ICA+SVM over conventional SVM

analysis. An elaborate validation and evaluation using clinical data sets were necessary to ensure the performance of the method. As a continuation, Chai *et al.* (2010) [29] performed an effective quantification of brain normal tissues and pathologies simultaneously using ICA+SVM. Detailed investigation and evaluation of effective volumetric measurements of normal and lesion tissues using synthetic and real multispectral MR images was performed. Accuracy and reproducibility of CSF, GM, WM, and WML volume measurements were also evaluated. The ICA+SVM clearly extracted the normal tissues and white matter hyper intensity lesions with low intra- and inter-operator coefficient of variations. Inconsistency involved in over-complete ICA was significantly reduced through ICA based SVM. The experiments conducted provided evidence that the ICA+SVM method has potential in clinical analysis of MRI. Our thesis work selected this work as the method for immediate comparison, to justify the improvement in SVM based classification. More details are given in Chapter 3.

2.8.3 ICA extensions to address the local feature extraction issues

Another critical issue of applying ICA in MRI analysis is the poor performance in identifying local features. In biomedical applications, relatively few works were observed to address this issue. In the case of MRI analysis, application of conventional ICA itself is a new topic; still researchers cannot exploit its full potential in clinical analysis. No previous works were reported in the literature to address this issue for brain MRI analysis.

Bauer *et al.* (2003) [152] proposed a neural network model for the identification and classification of small malign and benign skin lesions from ALA-induced fluorescence images. A self-organizing feature map or Generative Topographic Mapping (GTM) was used to cluster images patches

according to their inherent local features, which then can be extracted with ICA. These components were used to distinguish skin cancer from benign lesions, achieving an average classification rate of 70%. Clustering the data in a pre-processing step using SOM or GTM improved the overall recognition rate considerably. But accuracy was found to be strongly dependent on the cluster size.

Recently, Han and Li (2011) [53], proposed a novel feature selection method: Multi-resolution Independent Component Analysis (MICA) for large-scale gene expression data. This method resolved the weak points of the widely used transform-based feature selection methods such as Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Nonnegative Matrix Factorization (NMF) by avoiding their global feature-selection mechanism. In addition to demonstrating the effectiveness of the MICA in meaningful biomarker discovery, they presented a MICA based SVM (MICASVM) and LDA (MICA-LDA) to attain high-performance classifications in low-dimensional spaces. Comprehensive experimental comparisons with nine state-of-the-art algorithms on six high-dimensional heterogeneous profiles demonstrated the superiority and stability of the new approach. MICA-SVM showed near clinical level sensitivities and specificities as well as strong performance stability over its peers in classification. In this thesis work, we modified MICA for MRI analysis, and evaluated the improvement in classification performance by FCM and SVM. More details are available in Chapter 4.

2.9 Summary

In this chapter, recent research papers which are relevant to the present thesis work are discussed. The chapter started with a brief overview of methods in MRI analysis, especially single image analysis. Different classification

approaches are explained with more emphasis on the state of the art algorithms. Survey on supervised and unsupervised segmentation algorithms recommended supervised approach as more efficient and reliable. FCM and SVM based methods were found to be more popular in unsupervised and supervised category. Effectiveness of supervised and unsupervised multispectral analysis in MR image analysis is reviewed with some recently reported works. Methodological and data specific challenges and their possible solutions were also summarized in this chapter. The advantage of spectral unmixing using ICA in brain tissue analysis is focused in the last section. Review of challenges in MRI analysis using conventional ICA motivated us to explore more studies in over-complete ICA issue and local feature analysis. In MRI analysis these issues have great significance, since soft tissues or several types of small abnormalities are to be extracted with high accuracy from a limited number of bands.

Chapter -3

ICA AND BRAIN MRI

3.1 Introduction

Independent component analysis (ICA) is a generic model which helps to find a linear representation of non-gaussian and mutually independent data [144]. The main objective of ICA is to find the underlying, original signals or processes, which usually provide important information that cannot be directly measured from the observed signals. As a blind statistical signal processing technique, ICA can be applied to areas such as blind separation of mixed voices or images, analysis of several types of hidden data, speech and image recognition, data communication, sensor signal processing, biomedical signal processing and so forth. The basic concepts, algorithms and implementations are discussed in the first few sections of this chapter. After that, application of ICA in multispectral MRI analysis is discussed, and the common issues in ICA for MRI are illustrated with some recently proposed methods in the literature.

3.2 Basic concepts of Independent Component Analysis (ICA)

ICA can be considered as a method to solve the Blind Source Separation (BSS) problem, where original sources can be estimated from measured or observed mixture of signals, without any explicit knowledge of mixing system or signal sources. For example, the cocktail party problem [145], in which a number of people are talking simultaneously in a room and one is trying to extract the voice of a particular person from the rest of the speakers. ICA uses statistical independence of the source signals to solve these kinds of

blind signal separation problems. Algorithms that could successfully perform linear ICA was first introduced by Herault and Jutten in 1986 [153], and it has been established as a fundamental way of analyzing multivariate data in 1990's [142, 154].

Conventional ICA provides a linear decomposition of the data, like the more classical methods of multivariate analysis techniques, factor analysis and Principal Component Analysis (PCA). However, these classical methods sometimes failed to unmix or separate the underlying sources, mixed in the observed data. ICA uses the non-Gaussian structure of the data to recover the independent components that created the data. It is an unsupervised method, because it is not necessary to know the expected results of the system and it divides the measurements into different components without designing different experimental conditions or supervised training.

The basic idea of ICA is illustrated with the help of original source signals (source1, source2, source3, and source4), mixed signals (mixed1, mixed2, mixed3, and mixed4) and unmixed signals (extracted1, extracted2, extracted3, and extracted4) as shown in Fig. 3.1.

3.2.1 Definition

Let '**x**' be the random vector whose elements are the mixtures x_1, ..., x_n, and '**s**' be the random vector with elements s_1, ... , s_n. Let '**A**' be the matrix with elements a_{ij}. All vectors are considered as column vectors. Using the vector-matrix notation, the above mixing model is written as

$$\mathbf{x} = \mathbf{As} \tag{3.1}$$

Then, after estimating the matrix '**A**', its inverse can be computed, say '**W**', and we obtain the independent components simply by,

$$\mathbf{s} = \mathbf{Wx} \tag{3.2}$$

The statistical model in Eq. (3.2) is called independent component analysis, or ICA model, where it is assumed that components s_i are statistically independent, and unknown mixing matrix '**A**' is square**.**

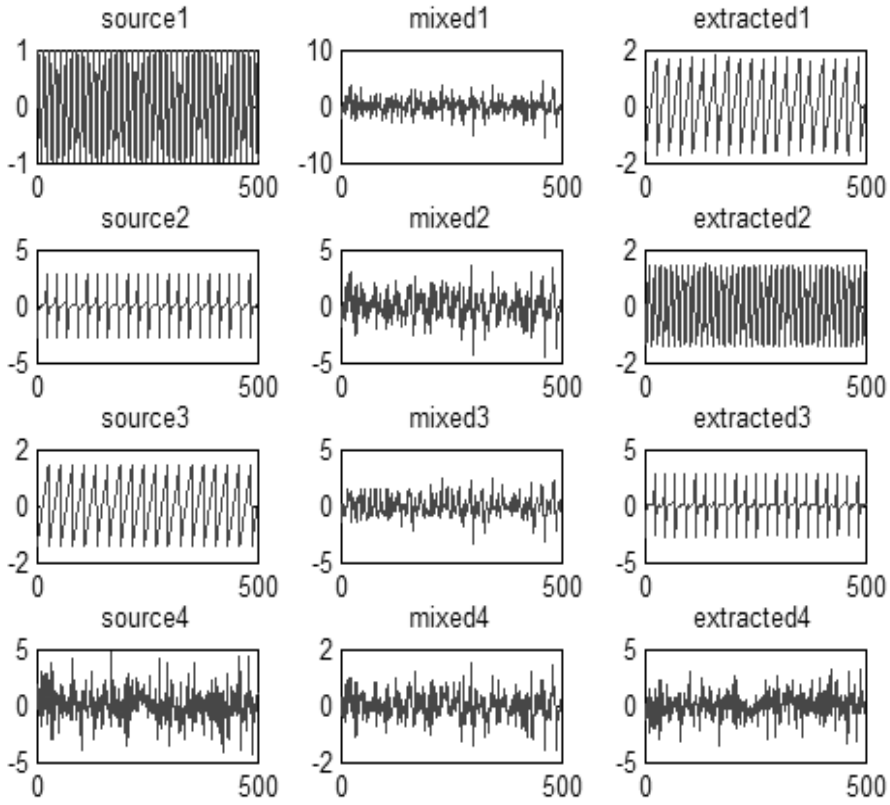


Fig. 3.1 The basic concept of ICA. First column represents the original source signals. From the four mixed signals (middle column), ICA extract the original source as shown in the last column.

3.2.2 Non-Gaussianity and Independent Components (ICs)

The key foundation of ICA is the assumption of non-Gaussianity of the independent components [142], which can be described as follows.

Two different random variables y_1 and y_2 are said to be independent, if information on the value of y_1 does not give any information on the value of

y_2, and vice versa. If the components of the mixed signals are created by physically separate and non-interacting source entities, they can be considered as statistically independent. i.e., s_1, ... , s_n in section 3.2.1 are independent, whereas mixtures x_1, ..., x_n are not. Theoretically, independence is defined in terms of joint and marginal probability density functions as described below.

Let $P(y_1, y_2)$ be the joint probability density function (pdf) of y_1 and y_2, then they are said to be independent if

$$P(y_1, y_2) = P_1(y_1) \,.\, P_2(y_2) \tag{3.3}$$

where $P_1(y_1)$ and $P_2(y_2)$ are the marginal pdf of y_1 and y_2 respectively, defined as,

$$P_1(y_1) = \int_{y2} P(y_1, y_2)\, dy_2 \tag{3.4}$$

The random source vector, **s**, in section 3.2.1, with multivariate density function P(**s**) is said to have statistically independent components if

$$P(\mathbf{s}) = \prod_{i-1}^{n} P_i(s_i) \tag{3.5}$$

Independence is a stronger property than uncorrelatedness[144]. According to statistical theory, if the variables are independent they are uncorrelated, but the converse is not true. PCA or factor analysis gives components that are uncorrelated, but uncorrelatedness is not enough to separate the hidden source components. Gaussian variables make ICA impossible because joint pdf of source variables are symmetric, from which we cannot estimate mixing matrix **A**. i.e., symmetric joint pdf does not contain any information on the columns of **A**. Projection of observed data in the direction of independence cannot be achieved through estimate of **A**.

In probability theory, the Central Limit Theorem (CLT) [144] states that the distribution of a sum of independent random variables tends toward a

gaussian distribution, under certain conditions. i.e., a sum of two independent random variables usually has a distribution that is closer to gaussian than the distributions of the original random variables.

In ICA model described in eq. (3.2), observed data signal x_i can be expressed in terms of source components s_k as

$$x_i = \sum_{k=1}^{n} a_{ik} s_k \,, \tag{3.6}$$

where a_{ik}, k=1, ... ,n are the constant coefficients.

According to CLT x_i's are more gaussian than s_k's. If $\boldsymbol{w}$ were one of the rows of the inverse of **A** in eq. (3.1), one of the independent components '**y**' can be estimated as a linear combination of **x**.

$$\mathbf{y} = \boldsymbol{w}^T \mathbf{x} \tag{3.7}$$

i.e., $$\mathbf{y} = \boldsymbol{w}^T \mathbf{A}\mathbf{s}. \tag{3.8}$$

where $\boldsymbol{w}$ is a vector to be determined. In practice, we cannot calculate $\boldsymbol{w}$ exactly, because we have no knowledge of matrix **A**.

Consider the transformation $z = \mathbf{A}^T \boldsymbol{w}$, then y can be expressed in terms of z as,

$$\mathbf{y} = z^T \mathbf{s} \tag{3.9}$$

$z^T\mathbf{s}$ is more gaussian than any of the s_i, and least gaussian when it equals to one of the s_i. This happens when only one of the elements, z_i of z is nonzero.

3.2.3 Objective functions

Estimation of the weights, **W** in ICA is based on contrast (or objective) functions that are calculated using some statistical properties of the data. The optimization (minimization or maximization) of these functions, and the relative adaptive change of the weights helps to estimate the final matrix **W**. Several objective functions based on different measures such as likelihood, entropy, and mutual information or more frequently on the approximation of

these statistical properties have been proposed for the estimation of the projection matrix **W** in the literature [142-144]. There are different ICA estimation approaches as explained below.

3.2.3.1 Measures of nongaussianity

a) Kurtosis: Higher order statistics uses kurtosis, the fourth order moment of random data as the classical measure of non-Gaussianity. Kurtosis of a random variable *x* is denoted as *kurt(x)* and defined by,

$$kurt(x) = E(x^4) - 3\{E(x^2)\}^2 \qquad (3.10)$$

Assume that *x* is centered (zero-mean) and has variance equal to one, then eq. (3.10) can be simplified to

$$kurt(x) = E(x^4) - 3 \qquad (3.11)$$

If *x* is a gaussian random variable *kurt(x)* is zero, since $E(x^4) = 3\{E(x^2)\}^2$. Kurtosis can be positive or negative; if *kurt(x)* is negative, then *x* is called subGaussian. A random variable with positive kurtosis is called superGaussian. Kurtosis, or its absolute value, has been widely used as the measure of nongaussianity in ICA, because of both computational and theoretical simplicity in it. As the value of kurtosis goes away from zero, distribution of the variable also becomes more non-Gaussian. However, it is very sensitive to outliers and are largely unaffected by structure in the middle of the distribution

b) Negentropy: It is another important measure of nongaussianity based on the information-theoretic quantity of (differential) entropy. The entropy of a random variable gives the degree of information provided by the observation of the variable. The differential entropy H of a random variable *y* with a probability density function, f(y), is given by

$$H(y) = -\int f(y) \log f(y)\, dy \qquad (3.12)$$

Independent components correspond to directions in which the differential entropy of $\mathbf{W}^T\mathbf{x}$ is minimized. Since differential entropy is not invariant for scale transformations, it is modified to obtain a linearly invariant version, negentropy J as follows,

$$J(y) = H(y_{gauss}) - H(y) \tag{3.13}$$

where y_{gauss} is a Gaussian random variable of the same covariance matrix as that of y. Thus, negentropy (or differential entropy) is defined as the difference between the entropy of a Gaussian random variable with the same variance as the observed random variable, and the entropy of the observed variable. Negentropy is zero when the observed random variable is Gaussian and positive when the observed variable is non-Gaussian.

The advantage of using negentropy as a measure of nongaussianity lies in the fact that it is well justified by statistical theory. While considering the statistical properties, it can be an optimal estimator of nongaussianity. However, it is computationally very expensive. Estimation of negentropy requires an estimate (possibly nonparametric) of the probability distribution function [144]. Simpler approximations of negentropy are found to be more useful in nongaussianity measurement.

c) Approximations to Negentropy : The approximations to negentropy were based on the maximum-entropy principle. In general we obtain the following approximation [144]:

$$J(y) \approx \sum_{i-1}^{p} k_i[\, E\{G_i(y)\} - E\{G_i(v)\}\,]^2 \tag{3.14}$$

where k_i are some positive constants, and v is a Gaussian variable of zero mean and unit variance. The variable y is assumed to be of zero mean and unit variance, and the functions G_i are some non-quadratic functions. If there exists only one non-quadratic function G, the approximation becomes,

$$J(y) \propto [E\{G_i(y)\} - E\{G_i(v)\}]^2 \tag{3.15}$$

We can fine tune this approximation by wisely choosing G_i that does not grow too fast, which yields more robust estimators. Some useful choices of G are,

$$G_1(u) = (1/\alpha_1)\,\log\cosh\alpha_1 u \tag{3.16}$$

where $1 \leq \alpha_1 \leq 2$ is some suitable constant, and

$$G_2(u) = -\exp(-u^2/2) \tag{3.17}$$

Approximations of negentropy give a very good compromise between the properties of the two classical nongaussianity measures, kurtosis and negentropy; they are conceptually very simple, robust and fast to compute.

3.2.3.2 Minimization of mutual information

It is another approach for ICA estimation, in which mutual information is used to measure the dependency between two random variables. Using the concept of differential entropy, define the mutual information '**I**' between m (scalar) random variables, y_i, $i = 1...m$ as follows,

$$\mathbf{I}(y_1, y_2 \ldots y_m) = \sum_{i=1}^{m} H(y_i) - H(\mathbf{y}), \tag{3.18}$$

where $H(y_i)$ give the lengths of codes for the y_i when they are coded separately, and $H(\mathbf{y})$ gives the code length when **y** is coded as a random vector, i.e. all the components are coded in the same code. In terms of negentropy eq. (3.18) can be rewritten as,

$$\mathbf{I}(y_1, y_2 \ldots y_m) = C - \sum_{i=1}^{m} J(y_i) \tag{3.19}$$

where C is a constant that does not depend on **W**.

From eq. 3.19, it is very clear that calculation of an invertible transformation **W** that minimizes the mutual information is roughly equivalent to finding directions in which the negentropy is maximized.

3.2.3.3 Maximum Likelihood (ML) estimation

It is a very popular approach to estimating the ICA model which is closely connected to the maximum likelihood principle. Pham *et al.* (1992) [155] formulated the likelihood in the noise-free ICA model, and then estimated the model by a maximum likelihood method. Let $\mathbf{W}= (\mathbf{W}_1, ...,\mathbf{W}_n)^T$ denotes the matrix $\mathbf{A}^{-1}$, the log-likelihood takes the form [19],

$$L= \sum_{t=1}^{T} \sum_{i=1}^{n} \log f_i(\mathbf{W}_i^T x(t)) + T \log |\det \mathbf{W}| \qquad (3.20)$$

where 'f_i' are the density functions of the s_i (here assumed to be known), and the $x(t)$, $t = 1, ...,T$ are the realizations of $\mathbf{x}$. The term $\log |\det \mathbf{W}|$ in the likelihood comes from the classic rule for (linearly) transforming random variables and their densities [156].

InfoMax principle [144] is based on maximizing the output entropy (or information flow) of a neural network with non-linear outputs. Let $\mathbf{x}$ be the input to the neural network whose outputs are $\varphi_i(\mathbf{W_i}^T\mathbf{x})$, where the φ_i are some non-linear scalar functions, and the $\mathbf{W_i}$ are the weight vectors of the neurons. For well chosen values of φ_i, the framework,

$$L= H(\varphi_1(\mathbf{W}_1^T\mathbf{x}), \varphi_2(\mathbf{W_2}^T\mathbf{x}), ..., \varphi_n(\mathbf{W_n}^T\mathbf{x})) \qquad (3.21)$$

can estimate the ICA model.

Maximum likelihood and mutual information are equivalent, for all practical purposes. The problem with maximum likelihood estimation is that the probability density function must be estimated correctly. Otherwise ML estimation will give completely wrong results. This problem does not occur in ICA estimation with a reasonable measure of nongaussianity. In many cases, we have enough prior knowledge on the independent components, and we don't need to estimate their nature from the data.

3.2.4 ICA implementation methods

Some of the ICA algorithms require preprocessing such as centering and whitening [144], of observed data and some may not. The most popular and widely referenced techniques for implementing ICA are discussed in the following sections.

3.2.4.1 Non-Gaussianity through kurtosis: FastICA

FastICA is one of the most widely referenced ICA techniques in the literature introduced by Hyvarinen and Oja (1997) [154]. FastICA attempts to unmix the inherent component characteristics from the given measured dataset based on the non-Gaussianity. In FastICA, a fast fixed-point iterative algorithm finds the projections that maximize the non-Gaussianity of components by their kurtosis as discussed in section 3.2.3.1

3.2.4.2 Non-Gaussianity through negentropy: InfoMax.

Bell–Sejnowski algorithm ([143], EEGLAB) attempted to extract non-Gaussian sources using negentropy, as discussed in section 3.2.3.1(b). It is a neural network gradient-based algorithm in which the learning rule is based on the principle of Information Maximization (*InfoMax*). The learning criterion is obtained by the maximum likelihood estimation of an ICA model, as discussed in section *3.2.3.3*. In general, as a neural network approach, it suffers from the problems inherent in gradient training and learning. The choice of the nonlinearities used also has a great role in the performance of the algorithm.

3.2.4.3 Joint Approximate Diagonalization of Eigenmatrices (JADE)

This approach is known as ICA by tensorial methods, since it uses higher-order cumulant tensors [145]. An Eigen value decomposition of the covariance matrix (second-order cumulant tensor) of the data, $\mathbf{C_X}$, transforms

the data such that the second-order correlations are zero. Similarly, fourth-order cumulant tensors make the fourth-order cumulants zero or as close to zero as possible. As the name implies, JADE involves the joint diagonalization of a number of matrices. i.e., it attempts to make all off-diagonals zero or close to zero as possible [157]. JADE is useful in low dimensional problems, but it has the limitation that it cannot be used with high dimensional data due to numerical reasons.

3.3 FASTICA

In this thesis work, FASTICA algorithm is used to implement ICA. The learning rule in Fast ICA finds a direction, i.e. a unit vector **W** such that the projection $\mathbf{W}^T\mathbf{x}$ maximizes the nongaussianity measured by the approximation of negentropy $J(\mathbf{W}^T\mathbf{x})$ as discussed in section 3.2.3.1. It is based on a fixed-point iteration scheme for finding a maximum of the nongaussianity of $\mathbf{W}^T\mathbf{x}$. An approximative Newton iteration [158] also can be used to find the maximum. It requires preprocessing techniques like centering and whitening to make the problem of ICA estimation simpler and better conditioned. The major steps in this computation are shown in Fig 3.2.

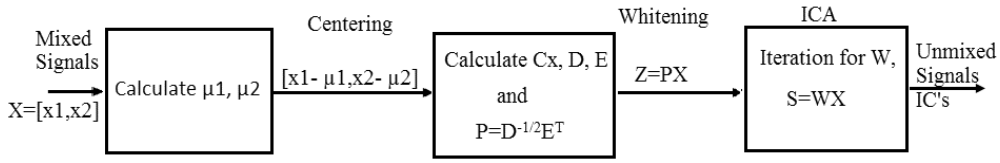


Fig. 3.2 FASTICA

3.3.1 Centering

Centre the observation vector **x** by subtracting its mean vector μ=E(x). It simplifies ICA algorithms. Thus the independent components are made zero mean, since

$$E(\mathbf{s}) = \mathbf{A}^{-1}E(\mathbf{x}) \qquad (3.22)$$

3.3.2 Whitening

Whitening or sphering is the process of linearly transforming the observation vector **x** by linearly multiplying it with some matrix **V** such that its components are uncorrelated and have unit variance. i.e.,

$$\mathbf{z}=\mathbf{V}\mathbf{x}, \tag{3.23}$$

where **z** is a new vector that is white. In terms of the covariance matrix, $E(\mathbf{z}\mathbf{z}^T) = \mathbf{I}$, where **I** is the unit matrix. A popular method for whitening is Eigen Value Decomposition (EVD) or PCA.

In PCA, an observed vector **x** is first centered by removing its sample mean. Then the vector is transformed by a linear transformation into a new vector, possibly of lower dimension, whose elements are uncorrelated with each other. The linear transformation is found by computing the eigenvalue decomposition of the covariance matrix. For a zero-mean vector **x**, with *n* elements, the covariance matrix **Cx** is given by:

$$\mathbf{C}\mathbf{x} = E(\mathbf{x}\mathbf{x}^T) = \mathbf{E}\mathbf{D}\mathbf{E}^T, \tag{3.24}$$

where $\mathbf{E} = (e_1, e_2 \ldots e_n)$, orthogonal matrix of eigenvectors of **Cx**, and

$\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, diagonal matrix of eigenvalue of **Cx.**

Whitening or sphering can be described as

$$\mathbf{z} = \mathbf{P}\mathbf{x} \tag{3.25}$$

where the whitening matrix **P** is defined as

$$\mathbf{P} = \mathbf{D}^{-1/2}\mathbf{E}^T \tag{3.26}$$

In the subsequent process of ICA estimation, iteration is performed on **z** instead of **x**. Methods like local PCA and random projection [144] can also be used to achieve the linear transformation and dimensionality reduction.

3.3.3 FastICA iteration for one Independent Component

A collection of whitened data, **z** is achieved through data pre-processing methods as described in the previous section. Let function, g() be the derivative of a non-quadratic function,

$f(u) = u^3$, or

$f(u) = -\exp(-u^2/2)$, or

$f(u) = \log(\cosh(u))$

Then the basic procedure of FASTICA to obtain one independent component can be summarized as follows.

Step1: Randomly choose an initial weight vector **W**

Step2: Let $\mathbf{W}^+ = E[\mathbf{z}\, g(\mathbf{W}^T\mathbf{x})] - E[g'(\mathbf{W}^T\mathbf{z})]\, \mathbf{W}$

Step3: Let $\mathbf{W} = \mathbf{W}^+ / \|\mathbf{W}^+\|$

Step4: if not converged, go back to *Step2*

The final vector **W** equals one of the columns of the (orthogonal) unmixing matrix. To estimate *n* independent components, this algorithm is to be repeated *n* times as discussed in the next section.

3.3.4 FastICA iteration for multiple units

Estimation of several independent components, can be achieved through repetition of one-unit FastICA algorithm using several units with weight vectors $W_1,...,W_n$. After each iteration, outputs, $W_1^T, ...,W_n^T$, are decorrelated to prevent the convergence of different vectors to the same maxima. Deflation scheme based decorrelation, symmetric decorrelation and iterative methods are the three commonly used approaches in the decorrelation process. FASTICA for multiple units incorporated these ideas also using additional steps as follows. Let C be the number of required independent components,

Algorithm:

for p = 1 to C,

step1: W_p = *Initial random vector of length N*

while W_p changes,

step2: Let $W_p^+ = E[\mathbf{z}\, g(W_p^T \mathbf{z})] - E[g'(W_p^T \mathbf{z})]\, W_p$

step3: $W_p^+ = W_p^+ - W_j W_j$

step4: $Wp = Wp^+ / \| Wp^+ \|$

end

end

Output : $\mathbf{W} = [W_1, W_2, \ldots W_C]^T$

$\mathbf{s=Wx}$

In the above algorithm, orthonormalization (decorrelation) is performed in *Step3* by deflation scheme approach. If we use symmetric orthonormalization, vectors are estimated in parallel, not one by one. In deflationary method, one by one estimation of vectors accumulates the estimation errors from the first vectors in the subsequent ones by orthonormalization [158]. Symmetric decorrelation can be accomplished by the update of **W** as,

$$\mathbf{W} = (\mathbf{W}\mathbf{W}^T)^{-1/2}\,\mathbf{W} \qquad (3.27)$$

where **W** is the matrix $(W_1, \ldots, W_n)^T$ of the vectors, and the inverse square root $(\mathbf{W}\mathbf{W}^T)^{-1/2}$ is obtained through EVD,

$$\mathbf{W}\mathbf{W}^T = \mathbf{E}\mathbf{D}\mathbf{E}^T \Longrightarrow (\mathbf{W}\mathbf{W}^T)^{-1/2} = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^T$$

3.3.5 Properties of FASTICA

Contrary to ordinary ICA algorithms based on (stochastic) gradient descent methods FASTICA has some desirable properties which makes it best

suitable to multispectral and hyperspectral approaches in remote sensing and medical image processing. Some of them are,

1. Fast convergence – As its name implied, the convergence is cubic (or at least quadratic) for FASTICA, where the convergence is only linear for stochastic gradient descent based approaches.
2. In contrast to the ordinary ICA algorithms, FASTICA is very easy to use since it has no step size parameters to choose.
3. FASTICA algorithm estimates independent components directly using a nonlinearity, g ( ), whereas many other algorithms choose the nonlinearity based on estimates of the probability distribution function.
4. One by one estimation of independent components in FASTICA is useful in exploratory data analysis, and decreases the computational load of the method in cases where a few independent components need to be extracted.
5. Like neural network based algorithms, FastICA is parallel, distributed, computationally simple, and utilizing less memory space.

All of these properties support FASTICA as a good choice in the implementation of ICA for higher dimensional applications. Performance of this method in the application of brain MRI for simultaneous analysis of multiple pulse sequences are discussed in the following sections.

3.4 ICA in brain MRI analysis

The greatest challenge in magnetic MR image analysis is the feature extraction from different pulse sequences, to be used in the pathological analysis for medical diagnosis. In this section, effect of ICA in feature extraction from multispectral brain MR images is analyzed with the help of some clinical and synthetic images. Also we discuss some major issues and

results of some case studies, using the algorithms addressing the problems in the literature. Recent researches on application of ICA in multispectral image analysis, as discussed in Chapter 2, demonstrated that multispectral and spatial information from brain MR images can be fully explored by statistical independency based transform, ICA [5, 33]. Based on these concepts, first, we have done a case study on clinical data to analyze the merits and demerits of this transform based feature extraction approach. Results observed for a multispectral cube formed by T1WI, T2WI and FLAIR images are shown in Fig. 3.3.

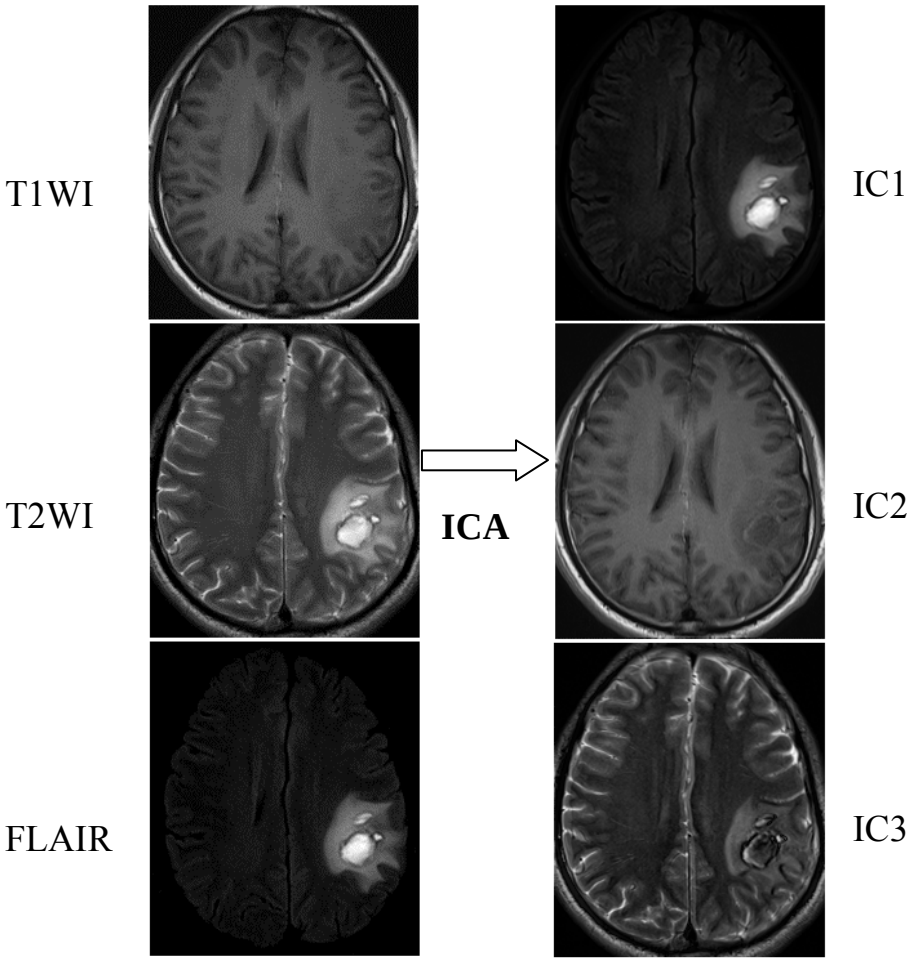


Fig. 3.3 ICA in brain MRI

First column shows the input images included in multispectral image cube, and the second column gives the results from ICA. From input images we can see that T1WI shows white matter information, but no abnormality details are available from it. T2WI locates tumor, surrounding edema and CSF. GM details are also visible in it. FLAIR images give the exact details of abnormalities, with traces of GM in the background. The input images can be considered as observed 2-d signals mixed by different brain tissues. Applying FASTICA on these images generates three different ICs, IC1, IC2, and IC3 as shown in Fig. 3.3. IC1 is similar to FLAIR image itself, whereas IC2 extract the details of WM with more information on abnormality locations affecting WM. IC3 gives information on CSF, but some traces of abnormalities mixed with CSF and GM are still visible in it. GM tissues are found to be very difficult to extract from these input images.

3.5 Issues in brain MRI analysis

ICA is found to be yielding very good results in brain tissue extraction, but constraints and limitations in ICA calculation often affects the accuracy of the clinical results, where soft tissues and small abnormalities are to be extracted with high accuracy from a limited number of bands. The major issues observed from the literature and case studies of multispectral MRI analysis can be summarized as follows.

1. ICA is a global transform based on global mean and covariance. Local details like small abnormalities were often neglected, while working with large amount of MR data [53].
2. Another issue occurs when number of tissues to be extracted is greater than the number of available images in multispectral suite (Over-complete ICA). ICA results in this case accumulate more than one source component in one of the ICs [5, 29].

3. Conventional classification systems and ICA use a simple linear approach. However, these linear models sometimes fail to represent complex data distributions found in practical applications [33].
4. Noise and other artifacts such as nonuniformity intensity, aliasing effects, patient movements in real MRI etc. adversely affects the efficiency of ICA in feature extraction [29].

In addition to these problem specific issues, following common ICA issues still exists [150],

5. Final sets of projection vectors produced by two different sets of random initial projection vectors (unit vectors) are generally different.
6. Due to the use of random initial projection vectors, the order that the ICs are generated is completely random, and does not necessarily indicate the significance or importance of an IC.

In recent studies, several attempts were made to address some of these issues so that ICA can be effectively used in multispectral MRI analysis. Major works in these directions are discussed with the help of some examples in the following sections.

3.6 Kernel ICA

Bach and Jordan, (2002) [148] presented an ICA extension to address the non-linearity issues in input data, known as Kernel ICA (KICA). The data in the input space $\mathbf{x} = (x_1, x_2 \ldots x_m)$ is mapped to a higher dimensional feature space $\boldsymbol{F}$ through a non-linear mapping Φ such that ,

$$\Phi: \mathbf{x} \in R^N \rightarrow \Phi(\mathbf{x}) \in \boldsymbol{F} \qquad (3.28)$$

and the nonlinear relation in the input data can be searched in the feature space $\boldsymbol{F}$ [147].

In the implementation, KICA does not calculate Φ, the implicit feature vector, explicitly. Instead of that, it computes kernel function K as inner

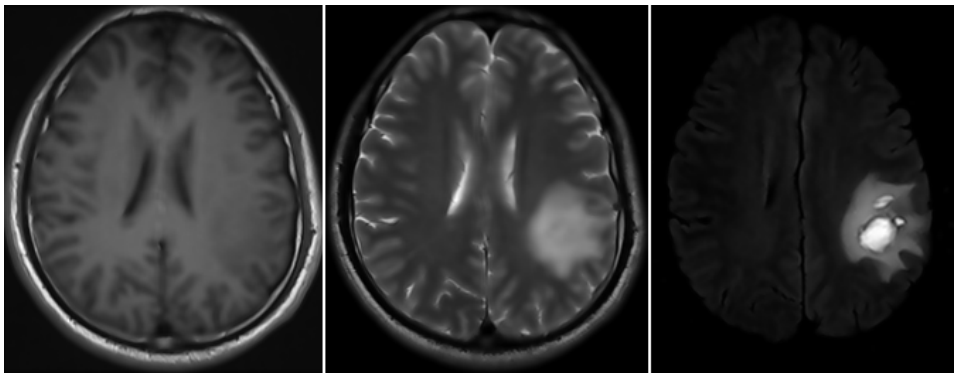
product of two vectors in $\boldsymbol{F}$ [147]. There are several kernel functions such as Gaussian kernel,

$$K(x_i, x_j) = \exp(-2\sigma^2||x_i - x_j||^2) \quad , \qquad (3.29)$$

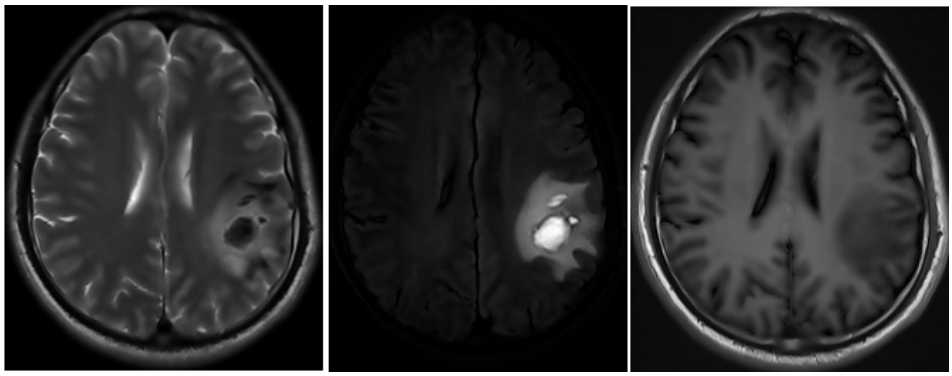
which corresponds to an infinite-dimensional Real Kernel Hilbert function (RKHS) space of smooth functions.

Application of KICA in a clinical brain MRI is shown in Fig 3.4. Tateyama et al., 2008 [147] demonstrated that KICA is a promising approach in brain tissue classification compared to conventional methods like neural networks, k-means clustering etc. However compared to ICA results, not much quality improvement observed for MRI analysis. Fig 3.4 (b) and Fig 3.4 (c) describes the difference in generated ICs by these two methods. On comparing the first column, it is seen that more specific extraction of CSF is possible with KICA. But, last column comparison shows that WM in ICA results is of better quality compared to that in KICA result.

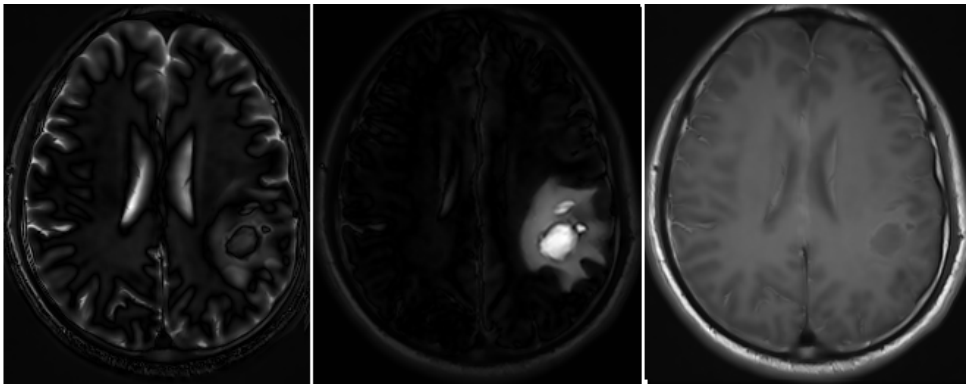
MRI analysis using KICA was found to be computationally very fast compared to conventional methods like neural networks [147]. However, the experiments with large sized MRI images were found to be too slow to converge, compared to ICA. On an average, more than half an hour difference observed between two methods for a 256x256 T1WI, T2WI and FLAIR image combination with windows7/Matlab implementation (running on a PC with Pentium Dual CPU/2.1GHz/4GB RAM).



(a) Input images



(b) ICA results



(c) Kernel ICA results

Fig. 3.4 Independent Components from ICA and Kernel ICA

3.7 Band expansion and Over-Complete ICA

Over-Complete ICA (OC-ICA) occurs when the number of observed signals is less than the number of sources to be unmixed. Suppose x_1, x_2, x_3 are the three observed mixtures, and we try to estimate 4 independent sources s_1, s_2, s_3, s_4 from the observed signals. Solving this problem becomes a linear system of equations as follows,

$$\begin{aligned} x_1 &= a_{11}s_1 + a_{12}s_2 + a_{13}s_3 + a_{14}s_4 \\ x_2 &= a_{21}s_1 + a_{22}s_2 + a_{23}s_3 + a_{24}s_4 \\ x_3 &= a_{31}s_1 + a_{32}s_2 + a_{33}s_3 + a_{34}s_4 \end{aligned} \qquad (3.30)$$

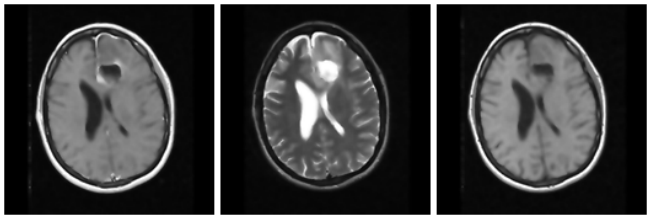
There exist many solutions to solve eq. (3.30), because the number of variables to be solved is greater than the number of equations. i.e., ICA must deal with an over-complete representation of a mixed model. Theoretically, the idea of the OC-ICA in MRI an analysis can be interpreted as pigeon-hole principle in discrete mathematics [150]. This issue occurs in a multispectral MR image analysis with three pulse sequences, when more than three brain tissues are to be extracted. In general, let L and p represent the number of different MRI sequences (holes) and number of brain substances (pigeons) to be classified respectively. $L < p$ implies that there are more pigeons than pigeon holes. In this case, at least one IC from ICA must accommodate more than one brain tissues as shown in Fig 3.5 (b). Ouyang et al. (2008) [5, 150] proposed a Band Generation Process (BGP) to resolve this issue, based on autocorrelation and cross-correlation as follows.

Let $\{Bi\}_{i=1}^{L}$ be the set of all original bands. Then

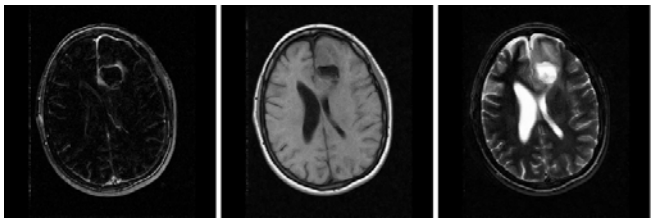
i. $\{B_i^2\}_{i=1}^{L}$ is the set of auto-correlated bands and

ii. $\{B_i B_j\}_{i=1,j=1,i\neq j}^{L}$ is the set of cross-correlated bands.

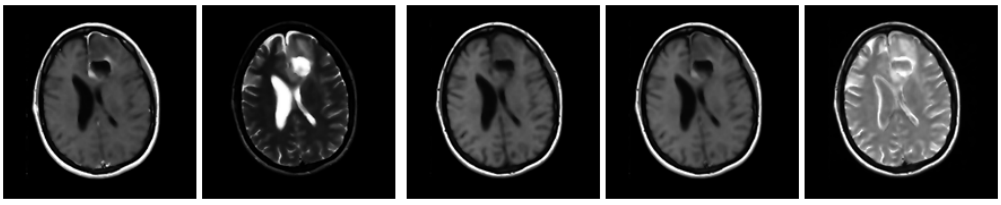
Results of band generation from clinical images (Fig 3.5a) are shown in Fig. 3.5 (c).



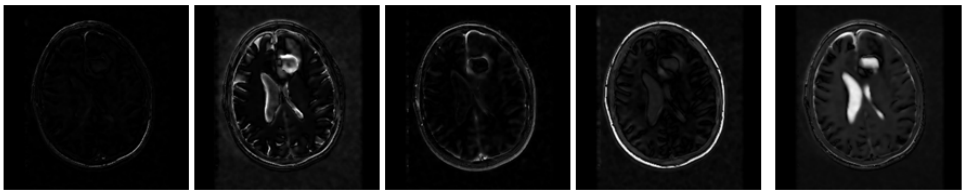
(a) Input images



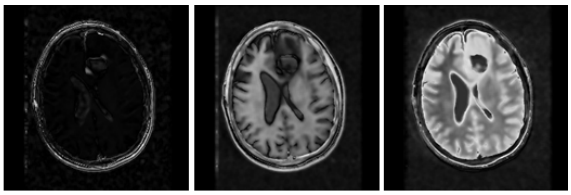
(b) ICA results



(c) Band expansion by autocorrelation and cross-correlation



Fat | CSF &Tumor | Muscles/skin | Connective Tissue | Lateral ventricle & Tumor



Skull | WM | GM

(d) Band expanded ICA results

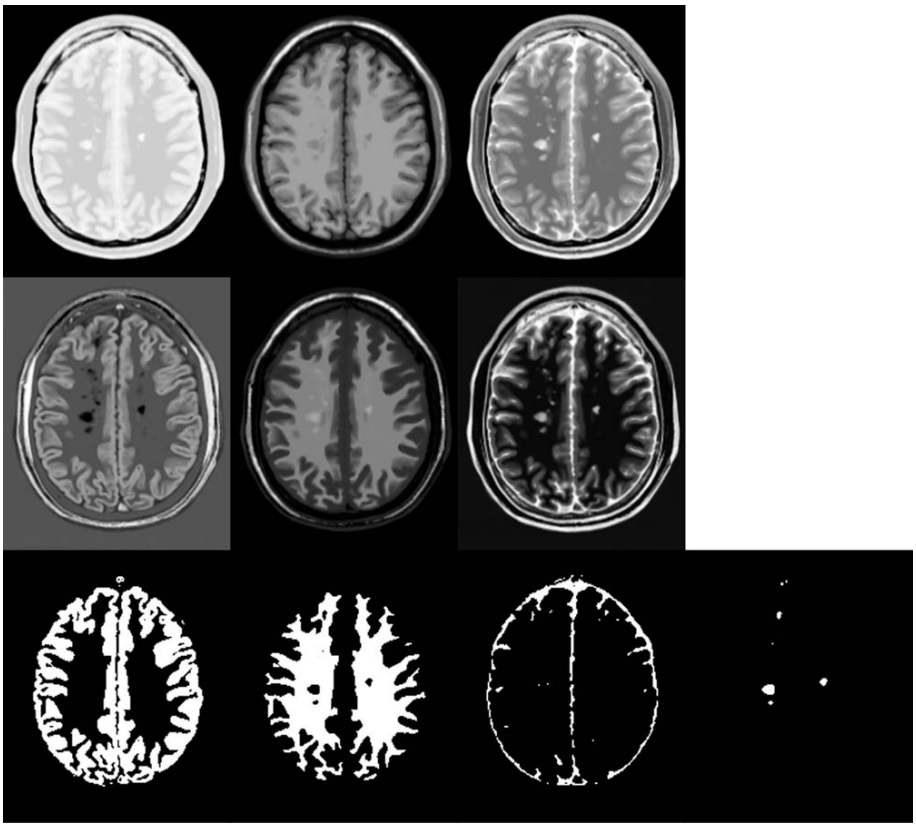
Fig. 3.5 Band expansion and ICA

ICA is then applied on band generated images and original input images as a combined multispectral data. Results are shown in 3.5(d). Soft tissues like GM, WM, CSF, muscles, fat etc. are extracted in different ICs, which make it easy to segment those brain tissues without any indexing.

These results support the findings by Y.C. Ouyang *et al.,* (2008) [150] that ICA alone cannot perform well, but their extensions and combinations with other band enhancement or feature analysis methods can significantly improve the performance of brain tissue classification [151].

3.8 ICA coupled with SVM for MRI analysis

A detailed quantitative analysis of ICA based SVM in clinical MRI is performed by Chai et al. in 2010. They demonstrated with the experimental results that SVM can provide better accuracy than other unsupervised methods or supervised methods, when more than one tissue features are accommodated in single independent component. Moreover, inconsistency associated with random initial projection vectors also can be avoided. Figure 3.6 shows the ICA results and classified brain tissues [29] for synthetic images from Brainweb. From Fig 3.6 (b) it is observed that, ICs contains extrameningial tissues and other major tissues. In ICA based SVM classification, they focused on brain tissues such as CSF, GM, WM and lesions. Best features from ICs were chosen with the help of an experienced radiologist, and extrameningeal tissues were removed from the classified results. The observed results are shown in Fig 3.6(c). SVM training and classification with optimal features from ICs helped to discriminate the relevant brain tissues from independent components. But in the extraction of lesion features it failed to identify the abnormality presence in some locations, which is reflected in the identification of diseased tissues in WM



Fig, 3.6 Independent components and classified brain tissues. Upper row is the synthetic input images with multiple sclerosis lesions (from left to right: PDI, T1WI, and T2WI). Middle row is three independent components. Lower row is GM, WM, CSF and white matter lesion images from SVM classification. [29].

(Figure 3.6 last row) also. Ambiguities observed in ICs affect the selection of best features; thereby it degrades the overall performance of the analysis. Inter operator variability and intra-operator variability in ICA based SVM classification for different brain tissues, measured by Chai et al. (2010) [29] indicates these issues as shown in Table 3.1.

Over-complete ICA issues and poor performance of ICA in small feature analysis lead to decreased values for performance measures like Tanimoto Index [29], sensitivity, specificity [77] and so forth. However, it is evident that feature selection from ICs which distinguishes the components with

better accuracy can lead to a more consistent and reliable classification with SVM in multispectral brain tissue analysis

Table 3.1 Inter and Intra operator variability: Mean Tanimoto Index of GM, and WM and Mean Sensitivity, Specificity, and Lesion Burden of WML Classification for Synthetic Image Data [29]

Operator		GM	WM	WML		
		TI	TI	Sens.	Spec.	Lesion Burden(mL)
Inter	Radiologist 1	0.823	0.899	0.870	0.999	2.36
	Radiologist 2	0.849	0.927	0.843	0.999	2.38
	Radiologist 2	0.827	0.914	0.859	0.999	2.40
	Mean	0.833	0.913	0.858	0.999	2.38
	Standard Deviation	0.014	0.014	0.013	0.00001	0.018
Intra	Measurement 1	0.823	0.899	0.870	0.999	2.36
	Measurement 2	0.841	0.921	0.868	0.999	2.39
	Measurement 2	0.828	0.922	0.868	0.999	2.41
	Mean	0.831	0.914	0.869	0.999	2.39
	Standard Deviation	0.009	0.013	0.001	0.00001	0.025

3.9 Summary

In the last few decades, ICA has become a standard tool in machine learning and signal processing applications. It is a very general-purpose multivariate analysis technique in which observed random data are linearly transformed into components that are maximally independent from each other. Due to the generality of the model, it finds applications in wide variety of fields. In earlier stages, assumptions of non-Gaussianity and independence discouraged the researchers to apply it in real-world applications. Computational load was another burden in multispectral and hyperspectral data analysis. However, later it was identified that applications dealing with scientific measurement devices give non-gaussian data. In addition to this, robust and efficient methods like FASTCA, InfoMax etc. were introduced to relax the assumption of independence and computational complexity.

Recent brain MRI studies using ICA and its extensions like kernel ICA, over-complete ICA and so forth have demonstrated the potential of ICA as a good feature extraction method for normal and abnormal tissue analysis. But it is observed that performance of ICA in classification of local features like small abnormalities is not so good. In the following chapters, we discuss our thesis work, the proposed ICA extensions for better feature extraction, and effect of the extended and hybrid approaches in brain tissue classification with the help of supervised and unsupervised classification techniques.

Chapter-4

PROPOSED ICA EXTENSIONS

4.1 Introduction

From the literature, we observed that a few extensions of ICA with efficient algorithms using wavelets, kernel functions etc. can significantly improve the classification performance. In our thesis work, we focus on the approaches to resolve the shortcomings of applications of ICA to brain MRI analysis, especially the inefficiency of ICA in identifying small abnormalities, and problems due to over-completeness. We propose the following ICA extension methods as a part of this thesis work to improve the performance of brain tissue classification from multispectral MRI.

1. A new Spectral Clustering ICA (SC-ICA) algorithm: An object based approach using spectral screening to retain local and global information with equal priority in brain tissue analysis.
2. Modified Multi-resolution ICA (MICA) algorithm for brain MRI analysis: Multiresolution analysis suppresses the global features and improves the priority of the local features, prior to ICA computation.
3. A new Multi-signal Wavelet ICA (MW-ICA) for brain tissue analysis: Band expansion with reconstructed brain images from detail coefficients helps to retain small abnormalities, thereby improve the classification performance also.

In this chapter, we elaborate these three approaches with relevant theory, algorithm, and performance comparisons with conventional ICA.

Experimental results and analysis of these algorithms in classification are included in Chapter 5. A brief discussion on the computational complexity of each method, and time required for typical cases are also added at the end of this chapter.

4.2 New Spectral Clustering ICA (SC-ICA)

In this method, a new spectral clustering algorithm based on Spectral Angle Mapping (SAM) [159] is used to extend ICA for improved feature extraction from multispectral MRI. SC-ICA provides equal priority to global and local features; thereby it tries to resolve the inefficiency of conventional approaches in abnormal tissue extraction. As a spectral angle based pre-processing technique, it can exploit intrinsic spectral characteristics of input signals and it is robust to input energy differences [159]. As a first step, the proposed method divides input multispectral MRI into different clusters by a spectral distance based clustering. Then, ICA is applied on the different objects grouped by spectral clustering.

4.2.1 Spectral Angle Mapping (SAM)

SAM is a technique that measures the similarity between the spectral signatures of objects in multispectral/hyper-spectral image cube. A multispectral image cube of '*n*' bands can be considered as '*n*' two dimensional matrices with each one corresponding to one band. A ray casting through each pixel in the multispectral image cube with bands B_1, B_2, …B_n, as shown in Fig. 4.1, results in a vector called the pixel vector, which gives the spectral signature of the tissue (or material) represented by that pixel.

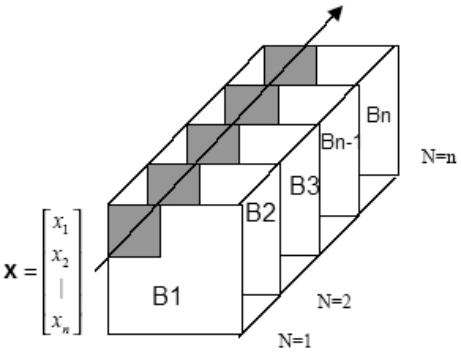


Fig. 4.1 A multispectral image cube and spectral signature

If 'x' and 'y' are two *n*-dimensional pixel vectors, where '*n*' is the number of bands in the input image, then the spectral angle between two vectors is given as [159, 160],

$$\alpha = \cos^{-1}\left(\frac{x.y}{\| x \| . \| y \|}\right) \qquad (4.1)$$

A two band spectral classification of spectral signatures $x=[x_1, x_2]$ and $y=[y_1, y_2]$ into $class_1$ and $class_2$, based on threshold 'α' using eq. (4.1) is shown in Fig. 4.2.

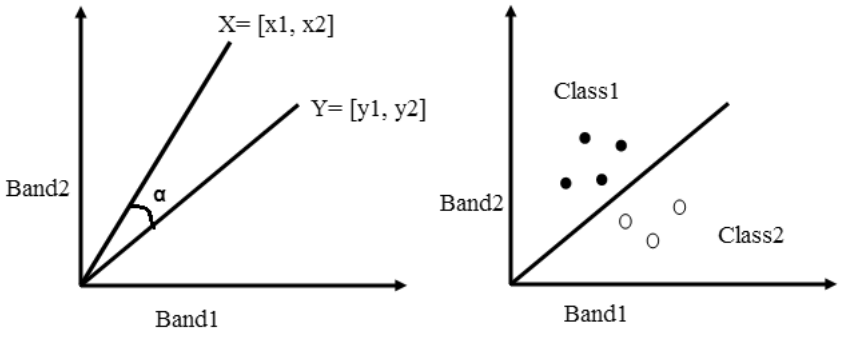


Fig. 4.2 Spectral Angle for two bands, X and Y are two 2-band pixel vectors

The spectral signatures can then be separated from one another as shown in the right side of Fig 4.2. It happens when there exists a sufficient difference in their angles; i.e. if the angle between two pixel vectors, α, exceeds a

particular threshold, α_{thr}. Larger α means 'x' and 'y' are very dissimilar, and easy to be differentiated. For example, spectral distances among different signatures of brain tissues such as GM, WM, and Lesion are tabulated in Table 4.1.

Table 4.1 SAM-based spectral distances between Tissue signatures

		WM	GM	Lesion
Synthetic	WM	0	0.24	0.03
	GM	0.24	0	0.12
	Lesion	0.03	0.12	0
Real	WM	0	0.2	0.05
	GM	0.2	0	0.16
	Lesion	0.05	0.16	0

Depending on the type of pulse sequences in the multispectral cube, spectral signatures also vary. In the example shown in Table 4.1, synthetic multispectral cube was formed from T1WI, T2WI, and PDI, whereas real multispectral data considered MRI pulse sequences T1WI, T2WI, and FLAIR images. Threshold value selection should be done according to the minimum angle shown by different subjects to yield an efficient clustering, which will be discussed in the next sections. According to Table 4.1, a minimum threshold, 0.03, can be selected for synthetic case, while selecting 0.05 or more for real cases.

4.2.2 Proposed SC-ICA

The main objective of this method is to retain less frequently occurred objects like small lesions, while dealing with massive amount of information in multispectral analysis.

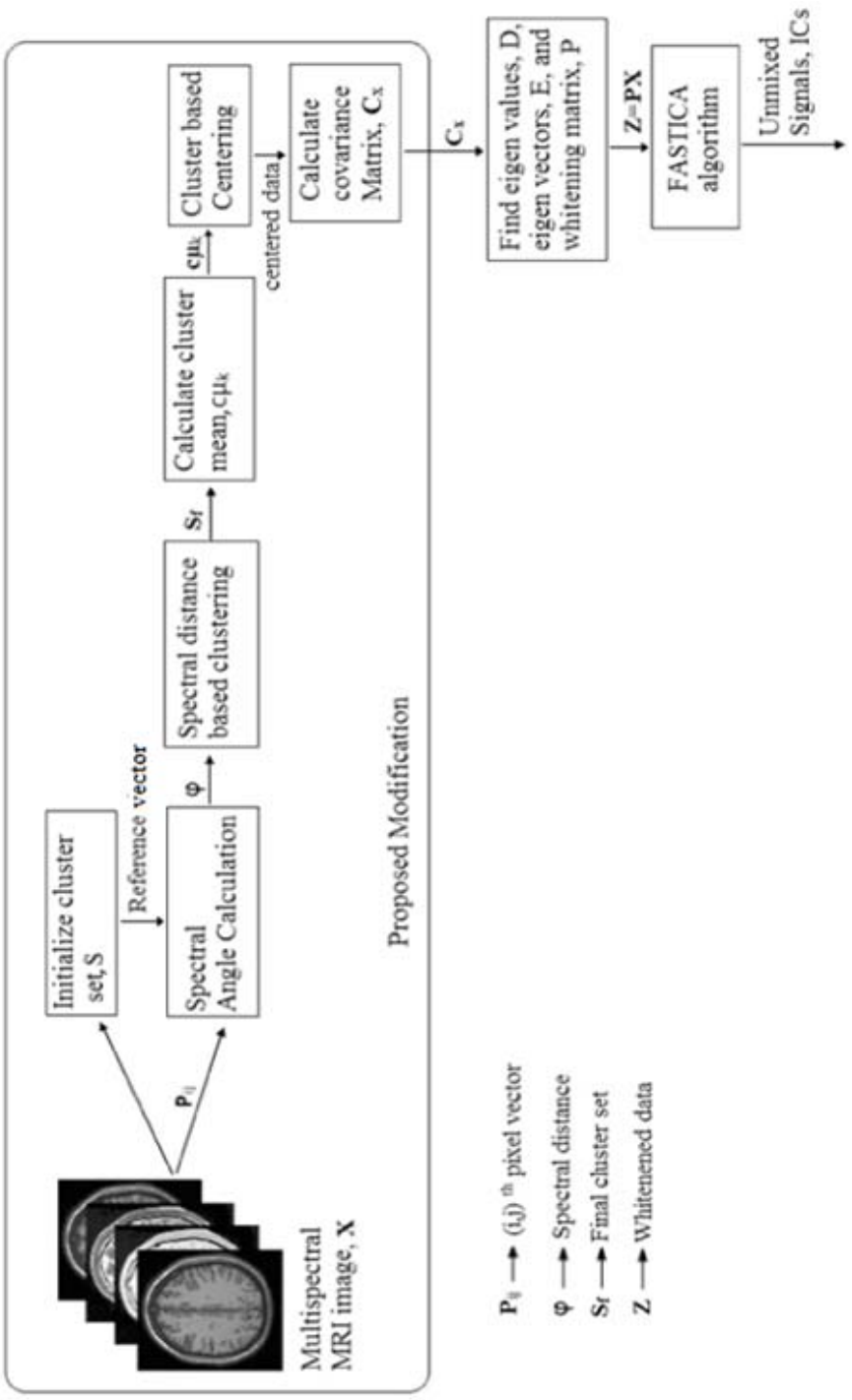


Fig. 4.3 Methodology of the proposed algorithm

The proposed method is divided into two parts; First part explains the method for clustering of multispectral MRI data to get object based signature groups, and the second part deals with centering and whitening part of the ICA algorithm based on statistical measures from clustered multispectral data. The proposed algorithm is outlined in Fig. 4.3. The grouped portion explains the details of proposed extensions to simple ICA process. Initially, input multispectral MRI cube is formed by placing registered T1WI, T2WI and PD/FLAIR image in 1st, 2nd and 3rd dimensions respectively.

***Input*:** Multispectral MRI cube, **X**, with pixel vectors $\mathbf{P}_{ij}$ and spectral angle threshold, φ_{thr}.

***Step1*:** Initialize unique cluster set **S** with a set containing first pixel vector in the image cube.

Step2*: *loop1 - for each $\mathbf{P}_{ij}$ do,

begin

loop2 -for each unique cluster c_k from **S**

begin

Calculate reference spectral signature,

$\mathbf{c\mu}_k$ = average pixel vector for c_k

Find φ_{ij} = Angle between $\mathbf{P}_{ij}$ and $\mathbf{c\mu}_k$.

if $\varphi_{ij} < \varphi_{thr}$, add $\mathbf{P}_{ij}$ to c_k . Continue with ***loop1***

else continue.

end

If $\mathbf{P}_{ij}$ is not added to any c_k in **S**,

Create a new cluster c_{k+1} in set **S.**

Add $\mathbf{P}_{ij}$ to $c_{k+1.}$

end

***Step3*:** Output unique cluster set $\mathbf{S_f}$ with elements c_1, c_2, c_3...

Algorithm 4.1 Part A: Spectral clustering

Input: Input dataset **X** and Cluster set $\mathbf{S_f}$ with elements c_1, c_2, c_3...

***Step1*:** For each cluster c_j in $\mathbf{S_f}$ do,

i. Find mean $\mathbf{c\mu_j}$

ii. Cluster based centering:

for each cluster elements (pixel vectors) $\mathbf{P_k}$ in cj,

Calculate, $\mathbf{P_k - c\mu_j}$

***Step2*:** Calculate covariance matrix $\mathbf{C_x}$ with the statistics computed from previous step.

***Step3*:** Compute **D,** diagonal matrix of $\mathbf{C_X}$'s Eigen values, and **E,** the orthogonal matrix of $\mathbf{C_X}$'s eigenvectors.

Calculate whitening matrix [23] **P** as follows**,**

$$\mathbf{P=D^{-1/2}E^T}$$

Step4*:** Whitening process: Use **P** from ***Step 3 to calculate whitened data,

$$\mathbf{Z=PX}.$$

***Step5*:** Apply FastICA algorithm on **Z** to obtain object specific unmixed independent components**.**

Output : Object based independent components in 2-D form

Algorithm 4.1 Part B: Spectral Clustering ICA (SC-ICA)

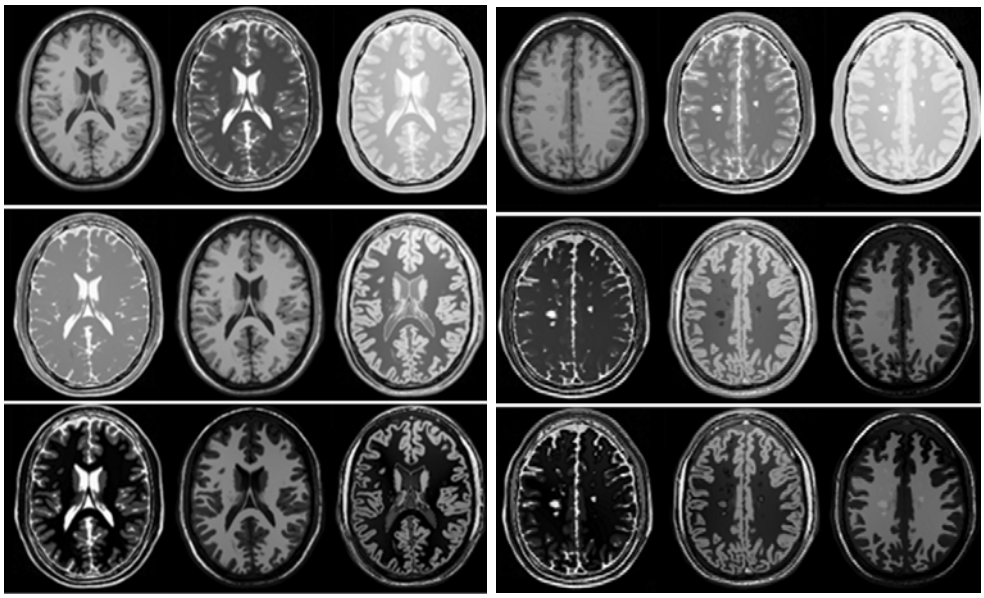
Let '**x**' be the L (here, L=3) dimensional pixel vector, which is linearly mixed by a set of m statistically independent tissues or tumor information, $s_1, s_2, \ldots, s_m$, by means of a Lxm mixing matrix '**A**'. Unknown signal sources, $s_1, s_2, \ldots, s_m$, in the image cube are to be separated using eq.(3.1) and eq.(3.2) as described in Chapter 3 Section 3.2. $\mathbf{P}_{ij}$ in Fig. 4.3 represents $(i, j)^{th}$ pixel vector. In this work, $P_{ij} = [x_1\ x_2\ x_3]^T$, where x_1, x_2, x_3 are $(i, j)^{th}$ intensity values from T1WI, T2WI and PD/FLAIR images respectively. First, cluster set, **S**, is initialized with first pixel vector. Calculation of spectral angle, φ, between each cluster mean (reference signature) and P_{ij} using eq. (4.1) classifies P_{ij} into appropriate group based on threshold value, φ_{thr}. Data centering of each cluster is separately done, and covariance matrix $\mathbf{C_x}$ is

calculated for whitening process**.** After that, optimizations and iterations of FASTICA are followed to generate ICs representing different brain tissues.

The entire procedure is summarized into two algorithms: Spectral clustering Algorithm in Algorithm 4.1 Part A and Spectral Clustering Independent Component Analysis (SC-ICA), in Algorithm 4.1 Part B.

4.2.3 Feature extraction using SC-ICA

Both synthetic and clinical datasets were given as input to SC-ICA as well as to ICA for feature extraction. Then, we analyzed the improvement in extracted features through a performance comparison of generated ICs from two methods. The same steps were repeated for SC-ICA on varying spectral angle threshold values also. ICs extracted by SC-ICA and ICA for 0% noise level synthetic sample data are available in Fig. 4.4. Left set (Fig. 4.4A) represent normal data and right set (Fig. 4.4B) gives results on abnormal images. Top row of both the sets represent input images, second row corresponds to ICs from conventional ICA and last row gives ICs from proposed method for threshold value 0.06. Independent components from SC-ICA for threshold value 0.06 (hereafter, referred to as SC-ICA_.06) were found to be yielding best results. It is evident from Fig. 4.4A and 4.4B that SC-ICA results are more specific compared to background dominating results from ICA. A clear picture of major tissues CSF, WM and GM is available from normal case ICs (Fig. 4.4A last row). In abnormal case analysis using SC-ICA, we could extract a unique feature set for each class, CSF, WM, GM and lesions (total 4 classes) from three components as shown in Fig. 4.4B last row.



A. Normal MRI slices B. Abnormal MRI slices

Fig. 4.4 Independent Components from synthetic images: Top row- Input images, Middle Row- ICs from ICA, Last Row- ICs from SC-ICA

Contribution of SC-ICA in identifying presence of abnormalities in other tissues, especially in the case of WM, was very evident without any indexing (Fig. 4.4B last row).

To study the variation in clinical cases on applying the new method, feature extraction from a real data as shown in Figure 4.5 is also added in this section. Sample slices of normal T1WI, T2WI and FLAIR images with the specifications as described in Chapter 1, are shown in Fig 4.5 upper row. IC's generated from ICA is given in the second row. To show the effect of more clusters in feature extraction, SC-ICA results are shown for threshold values 0.06 in 3rd row and for 0.03 in last row.

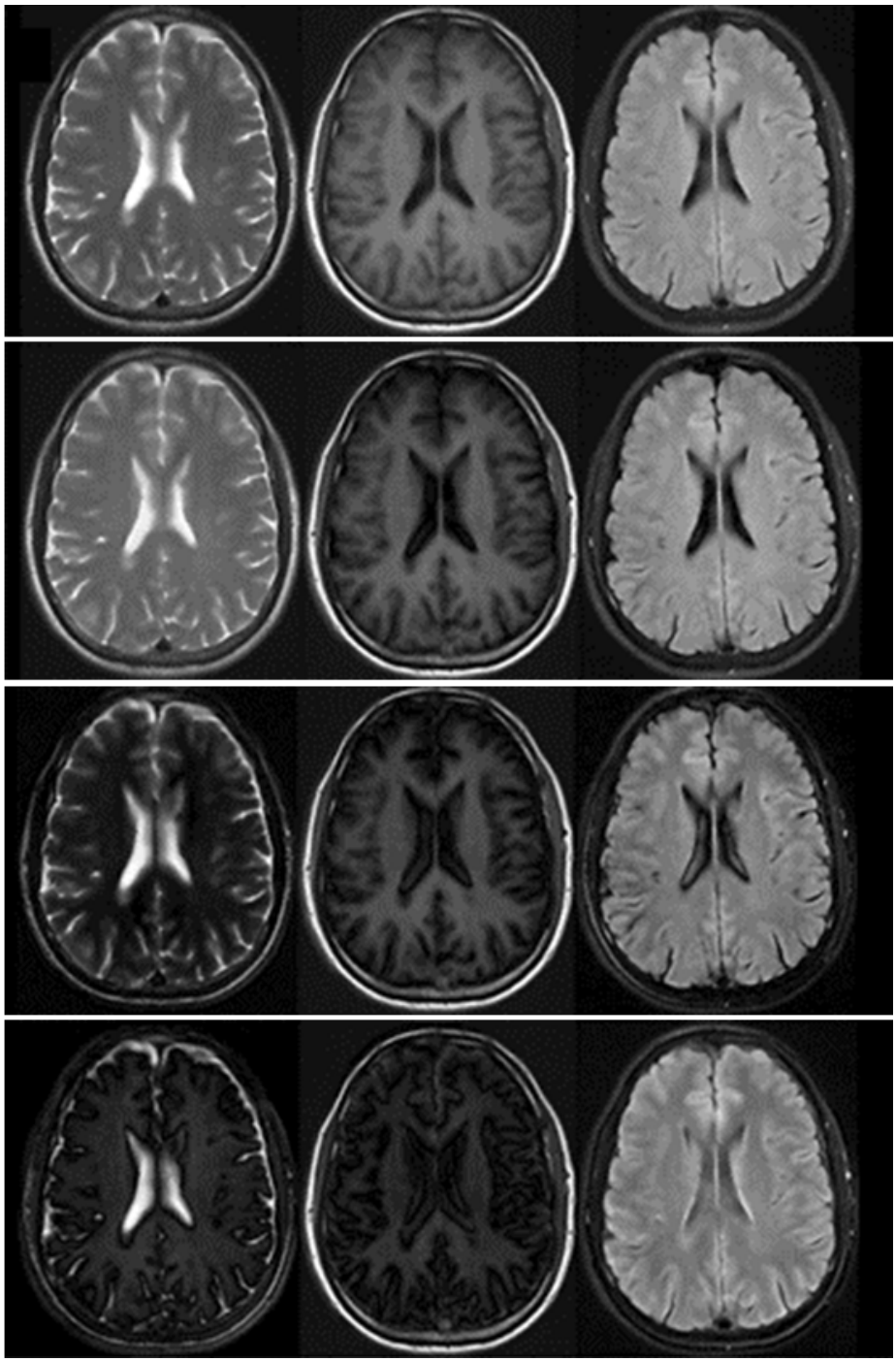


Fig. 4.5 Independent Components from normal real case: First row-Input images, 2nd row – ICs from ICA, 3rd row- ICs from SC-ICA_0.06, 4th row- ICs from SC-ICA_ 0.03

Fig. 4.5 shows that features available in SC-ICA results are more distinguishable compared to uncertain results from ICA (second row). On comparing second and third row, ICA and SC-ICA results, it is seen that a clear picture of major tissues CSF, WM and GM is available from 1st, 2nd and 3rd components of 3rd row. From the input images, it is seen that FLAIR image cannot contribute much to the multispectral analysis for tissue classification in normal case. Still maximum available information is unmixed using SC-ICA for GM, which is evident from Fig 4.5, 3rd and last row. Background domination is reduced in a considerable manner for SC-ICA_.03 also (last row in Fig. 4.5), compared to ICA results in clinical cases. But, it is observed that last row results failed to give the expected quality. From Table 4.1, it was noted that, spectral angle between lesion and WM gives the minimum value 0.05. In a normal case data, no lesion details are to be considered. Therefore, smaller threshold values will generate clusters for the same tissue group which will result in some distortions as shown in Fig 4.5 last row, and it will adversely affect the classification also. A detailed evaluation and analysis of more clinical cases with different threshold values in supervised and unsupervised MRI classification can reveal much detail on the potential of SC-ICA in brain tissue classification, which will be discussed in the next chapter.

4.3 Multiresolution ICA modified for MRI analysis

In SC-ICA, we observed that spectral clustering has improved the feature extraction from brain MRI, especially for the cases with the presence of small abnormalities. However, it is highly dependent on the spectral angle threshold selection. In this section, a wavelet based approach is discussed as an alternate to retain the priority of the small objects. Wavelet transform, feature selection and inverse wavelet transform from the recently introduced Multiresolution Independent Component Analysis (MICA) algorithm in

microarray classification [53] is proposed with relevant modifications for MRI analysis to improve the brain tissue classification. In this work, each 2-D MRI slice is reshaped into a 1-D signal, and used as a component signal in the multisignal form of multispectral data. Multisignal wavelet analysis is applied on these signals, and detail coefficients are modified to suppress the global features. Then, FASTICA is applied on the newly reconstructed signals to get relevant features.

4.3.1 Multisignal wavelet analysis

Wavelet transform of a signal is computed from projection of the signal onto the scaled and shifted version of the basic function, mother wavelet ψ(t) satisfying the condition,

$$C\psi = \int_{-\infty}^{\infty} \frac{|\psi(s)|^2}{|s|} ds < \infty \qquad (4.2)$$

which implies that wavelet has a zero average. i.e.,

$$\int_{-\infty}^{\infty} \psi(x)dx = 0 \qquad (4.3)$$

The mother wavelet, ψ(t), decomposes the signal into basis functions of the form,

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \qquad (4.4)$$

where a is the scaling factor and b is the translation factor.

Considering the binary partitions of the space, dyadic wavelets can be defined as [55]

$$\Psi_{k,i}(t) = 2^{\frac{-k}{2}} \psi\left(2^{-k}t - i\right), \qquad (4.5)$$

where 2^k represents the scaling factor and i is the translation factor. Multiresolution wavelet transform with dyadic wavelets provides a time-

scale domain representation of the signal under consideration, where time and scaling information can be studied simultaneously.

In the case of higher dimensional signals, wavelet analysis can be done in spatial or spectral direction. In the proposed method, discrete orthonormal bases of wavelets are computed using Mallat's algorithm for signal decomposition [50] extended to multidimensional signals, as shown in Fig. 4.6. Application of discrete wavelet transform in spectral direction will decompose each spectral signature into a set of composite bands that are linear, weighted combinations of the original spectral bands [55].

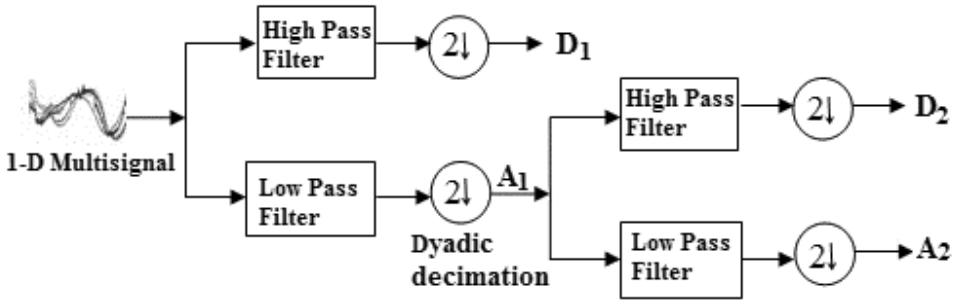


Fig. 4.6 Multisignal wavelet analysis of multispectral data

A low pass filter and its corresponding high pass filter are simultaneously applied on input signal at each level "i" in its column direction (spectral domain). Spectra are decomposed into approximation coefficient A_i and detail coefficient D_i as shown in Fig.4.6. Irrelevant elements involved in the signals are eliminated by a dyadic decimation, which reduces the original resolution into half of its length (downsampling). This procedure is recursively applied on approximation coefficients to give increasingly smoother versions of the original signals.

Mallat's algorithm works perfectly, when the number of input spectral bands is an integer power of two. Otherwise, spectral extensions techniques such as

mirroring the values of the signal with respect to the boundary, repeating of the boundary value, and zero padding [55] should be applied to the original spectral bands. Wavelet transforms can preserve both low frequency and high frequency features in a separate basis vectors, which is exploited in the feature selection of MICA. Reverse process of analysis is adopted in wavelet synthesis or reconstruction. The decomposed spectral data is lengthened by inserting zeros between every two samples of the signal (upsampling), and inverse filtering is applied. In short, wavelet decomposition involves filtering and downsampling, whereas wavelet reconstruction involves upsampling and filtering.

4.3.2 Existing multiresolution ICA for gene array classification

Recently proposed MICA method [53, 161] extracts contributions from almost all local features by taking low frequency subbands as such, and only the most important global features through elimination of high frequency details. Moreover, the redundant global feature suppressing mechanism in MICA acts as an automatic de-noising mechanism, since it modifies the coarse level coefficients so that unnecessary signals are filtered. The major steps included in the MICA based gene array analysis are shown in Fig. 4.7. A gene expression profile with p samples across n genes is given as input. MICA conducts an L-level discrete wavelet transform for each sample to obtain a sequence of detail coefficient matrices, and approximation coefficient matrix. The feature selection part in MICA keeps the most important detail coefficients with the help of PCA, and irrelevant details are suppressed. Inverse wavelet transform is applied on the updated coefficient matrices to construct the meta profile corresponding to input gene expression profile. ICA is then applied on this meta profile to obtain mixing matrix and

ICs. As a last step, each sample is decomposed in the subspace spanned by all independent components to preserve

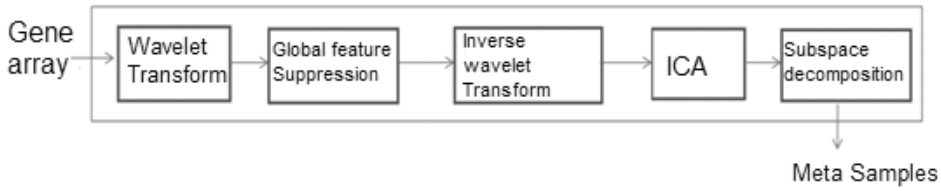


Fig. 4.7 MICA for microarray classification ([53])

the significant prototype of input data. SVM is applied on these meta samples to perform an efficient classification.

4.3.3 Proposed MICA for MRI analysis

In chapter 3, we discussed that conventional ICA, as a global transform, fails to extract the significant and precise local information from multispectral MRI data. Wavelet analysis, synthesis and global feature suppression part of MICA algorithm is considered in our proposed method, which attempts to preserve the local features with more priority than global features.

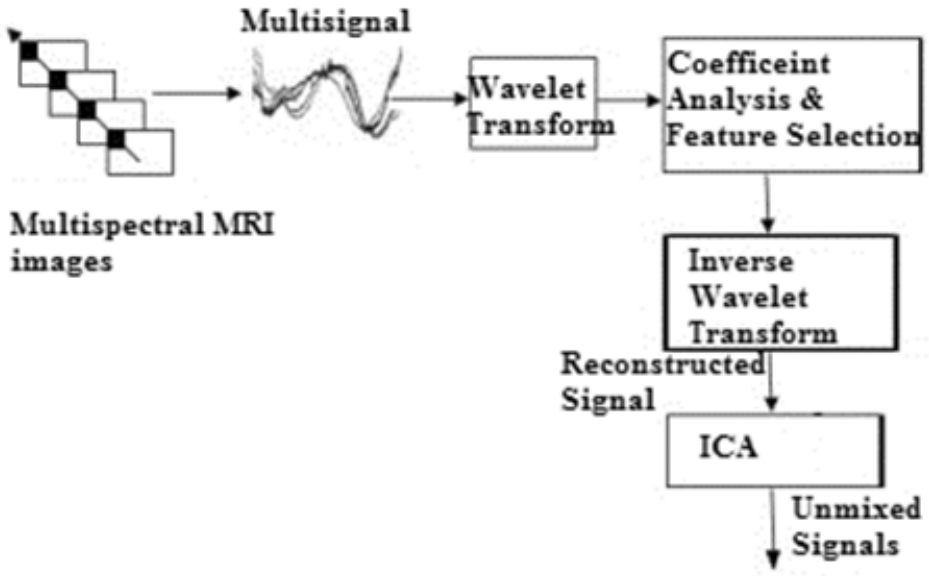


Fig. 4.8 Proposed MICA for MRI analysis

FASTICA method is applied on modified samples to generate ICs. Subspace decomposition of MICA algorithm has no use in MRI analysis. The proposed steps in MICA modification for MRI analysis are shown in Fig. 4.8, and main steps are explained in Algorithm 4.2.

4.3.3.1 Pre-processing and multisignals formation

MRI sequences for clinical trials are usually acquired with different size and orientation. Registration of the images [162] to form a co-registered multispectral suite is the preliminary step in the analysis process. Each pixel vector in a multispectral image acts as the spectral signature corresponding to that pixel, and a collection of those spectral signatures generate multisignals as shown in Fig. 4.8.

4.3.3.2 Wavelet analysis and modification of detail coefficients

1-D wavelet analysis of multisignals generates approximation coefficients and detail coefficients for different resolutions. Let $\mathbf{X}= [x_1, x_2,\ldots, x_p]^T$ be the spectral signature of a pixel as shown in Fig. 4.8, where '*p*' is the number of bands in multispectral image. *L*-level discrete wavelet decomposition of **X** forms a set $\{D_1, D_2...D_L, A_L\}$, where D_i's are the detail coefficient at level 'i' and A_L is the approximation coefficient.

Detail coefficients contain global features. So global feature selection is done by thresholding and recalculation of D_i's using a level threshold μ, where $1 \leq \mu \leq L-1$. Wavelet coefficient analysis and principal component calculations for different threshold values are performed as described in Algorithm 4.2.

Input: MRI images I1, I2, I3, I4 etc.

Step1: Pre-processing and Multisignals formation.

Step2: Wavelet Analysis and modification of detail coefficients, D_1, D_2...D_L

For a level threshold μ,

if $1 \leq j \leq \mu$,

Conduct principal component analysis of D_i's to get PC matrix **U** = [U_1, U_2 …U_p] and corresponding score matrix **S**= [S_1, S_2…S_p].

Reconstruct the original D_j by $D_j = (1/n_j)D_j In_j + S_1 x U_1^T$

if $j > \mu$,

Update each detail coefficients matrix D_j by using loading vectors U_1, U_2...U_k consisting of 100% explained variance percentage and their corresponding vectors in the score matrix such that

$$D_j=(1/n_j)D_j In_j In_j^T+[S_1,S_2,\ldots,S_k]x[U_1,U_2,\ldots U_k]^T ,$$

where n_j is the number of rows in D_j , and In_j is a unit vector of size n_jx1.

The new wavelet decomposed data at level '*L*' is given by

T* = {D_1, D_2...D_L, A_L}

Step3: Wavelet synthesis or reconstruction

Apply inverse discrete wavelet transform on **T*** to get the reconstructed signal **X***, i.e., **X* = IDWT (T*).**

Step 4: Apply FASTICA algorithm on X* to generate ICs

Output: Independent components with enhanced local features in 2-D form

Algorithm 4.2 Proposed MICA algorithm for MRI analysis

4.3.4 Proposed MICA and synthetic MRI

In order to explain the proposed method, synthetic images from Brainweb [20] database, containing information of MS lesions, are considered in this section. Fig. 4.9 top row shows T1WI, T2WI and PDI from left to right.

Implementation of step 2 in the proposed algorithm is performed with a 3-level wavelet decomposition of input multisignals, using Daubechies-4 (db4) [163] wavelet. Unlike gene array classification, MRI application has a limited number of bands (here, 3), which restricts the level of decomposition and threshold value to a few options. Explained variance [53] observed for threshold value 2, is given in Table 4.2.

Table 4.2 Explained variance for detail coefficients

level	PC1	PC2	PC3
1	99.9944	0.0056	0
2	99.9997	0.0003	0
3	100	0	0

Going to the deeper levels, it is observed that majority of the information is accumulated in first 1 or 2 components, which will reduce the dimensionality of the coefficients to be considered and computational overhead of the algorithm. In Fig 4.9, middle row shows the conventional ICA results, and last row represents the ICs from modified MICA. It is observed that abnormalities (circled portion) are so clear in MICA results that radiologists can visually analyze the points directly from the independent component itself, without performing a classification. Comparing first ICs from proposed method and ICA, it is evident that, in contrary to ICA result which accumulated CSF and lesion details in first component, proposed method separates them into different components. However, global feature suppression is found to be influencing the analysis of global brain tissues like WM and GM to a great extent as shown in last row, and it will be discussed in the next chapter as a part of the result analysis of proposed method for supervised and unsupervised classification.

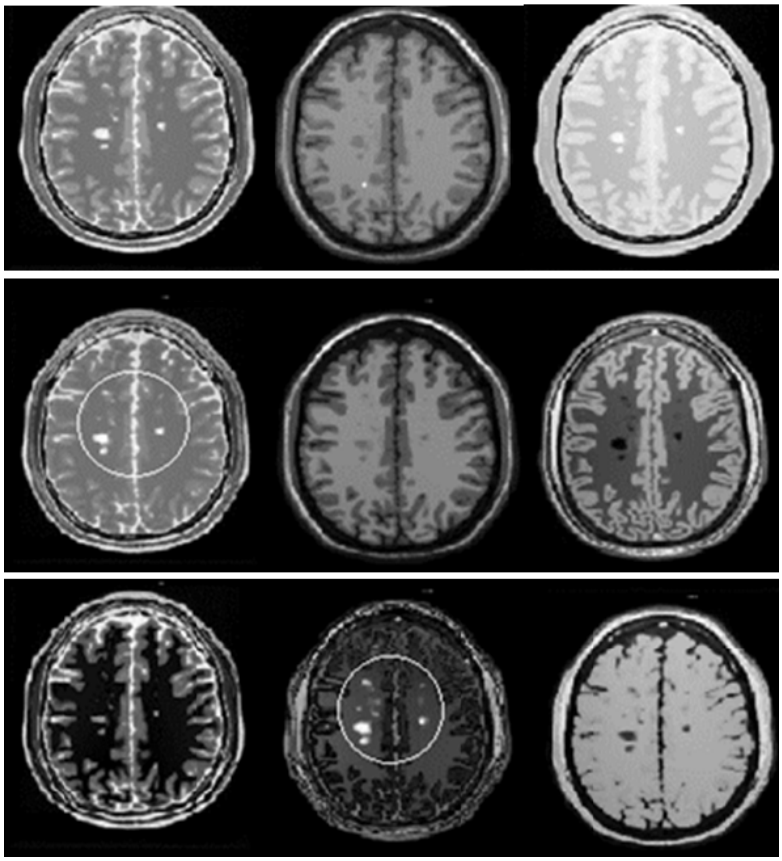


Fig. 4.9 Feature extraction from synthetic data using ICA and MICA, 1st row – Input images, 2nd row– ICs from ICA, 3rd row- ICs from modified MICA.

4.4 New Multisignal Wavelet ICA (MW-ICA)

We introduced SC-ICA as a method to resolve the inefficiency of ICA in small abnormalities detection, but threshold selection was observed as an issue in it. We modified MICA algorithm to apply in MRI analysis, exploiting the benefit of wavelet analysis in preserving local information. However, the global suppression is found to be adversely affecting the normal brain tissue classification. In this section, we discuss our third attempt, an improved source separation method, Multisignal Wavelet

Independent Component Analysis (MW-ICA) to provide an efficient, automated classification for multispectral MRI. First, local and global characteristics from brain MRI images were extracted by multisignal wavelet analysis. Then, detail coefficients were considered for wavelet reconstruction to enhance the original data with mutually independent details. Reconstructed signals were appended to original input signals to form an enhanced input signal set for ICA. As a band expansion method using wavelet, it can address two limitations of ICA in brain MRI analysis; inefficiency in identifying local features, and over-completeness problem. Wavelet decomposition of the spectra and ICA are the core concepts used in this algorithm. Major steps involved in this method are depicted in Fig. 4.10.

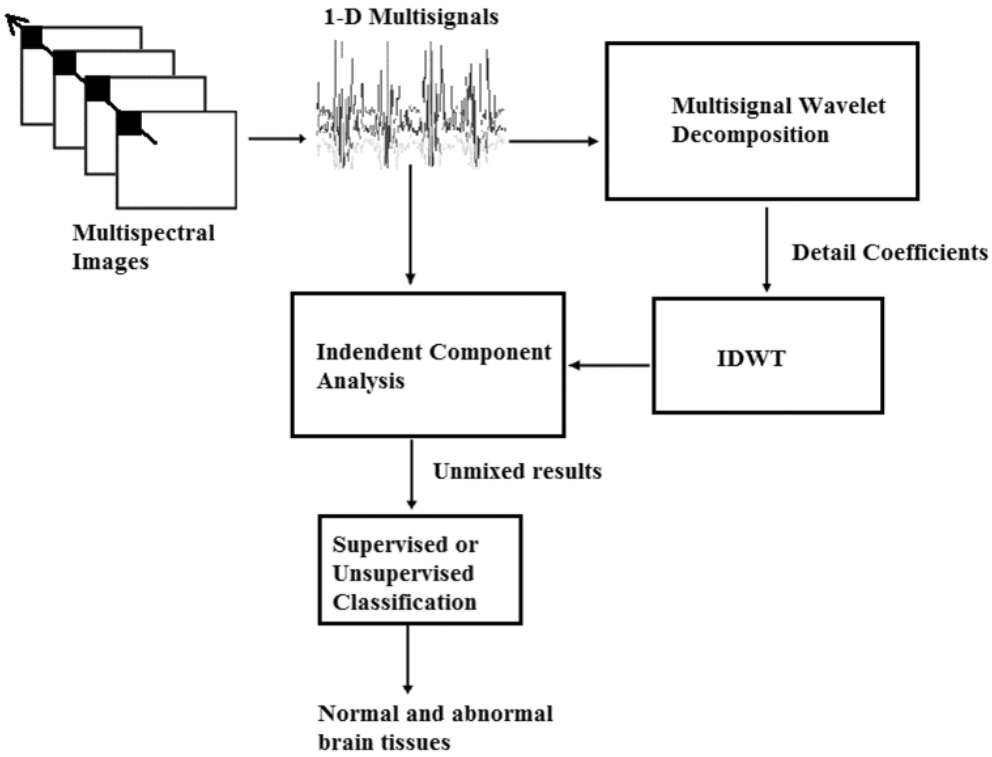


Fig. 4.10 Proposed MW-ICA based classification

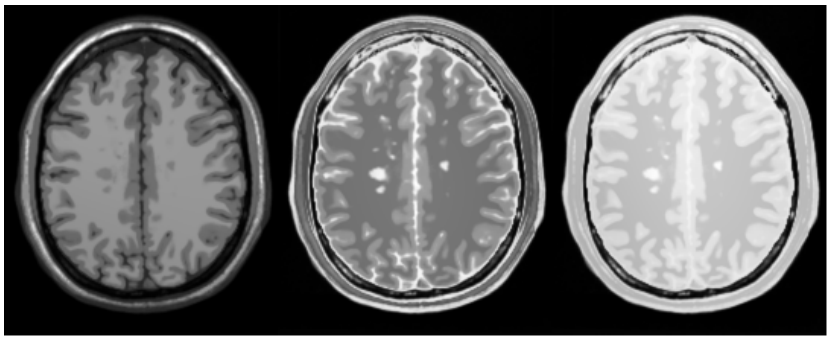
Multisignals are formed from multispectral MRI image cube, repeating the same procedure as discussed in Section 4.3. Wavelet analysis, synthesis and band expansion involved in the implementation is described in the following sections with the help of synthetic brain MR data.

4.4.1 Wavelet analysis and band expansion

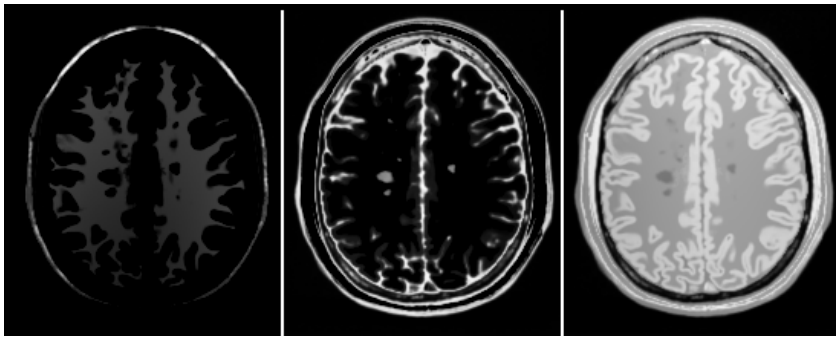
Co-registered MRI sequences, as shown in Fig. 4.10, are included in the input multispectral data. Each pixel vector represents the spectral signature of the area specified by that pixel. Consider input multispectral image as a collection of spectral signatures, represented by ray passing through the pixel vector (the shaded portion of multispectral image). Apply 1-D multisignal wavelet decomposition to these signals to divide the spectral domain into low frequency and high frequency components, as described in Section 4.3.

Wavelet transforms can preserve low frequency and high frequency features through multiresolution analysis, as explained in Section 4.3.1. Studies showed that high frequency subband signals are independent, and low frequency subbands are weakly dependent [55]. The proposed method targets an efficient classification, preserving both local and global features with same priority. So, independent high frequency subbands coefficients were considered in the multisignal wavelet reconstruction [50].

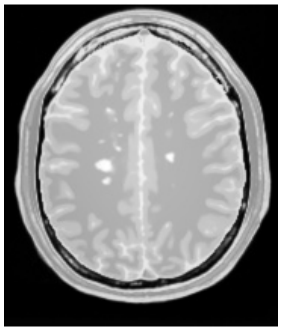
Level of decomposition plays an important role in multiresolution analysis. Decomposing the spectra into too many, smaller levels will result in subband signals with more local details in high pass components [55], making approximation coefficients deviating a lot from original signals. So the decomposition level in this work was restricted to the greatest integer contained in $\log_2(d)$, where "d" is the dimension of the multispectral cube. The reconstructed signals from detail coefficients were considered for band expansion, and appended to input multisignal dataset.



a. Input images



b. Reconstructed images from detail coefficients



c. Reconstructed image from approximation coefficients

Fig. 4.11 Multisignal wavelet analysis and synthesis of MR images

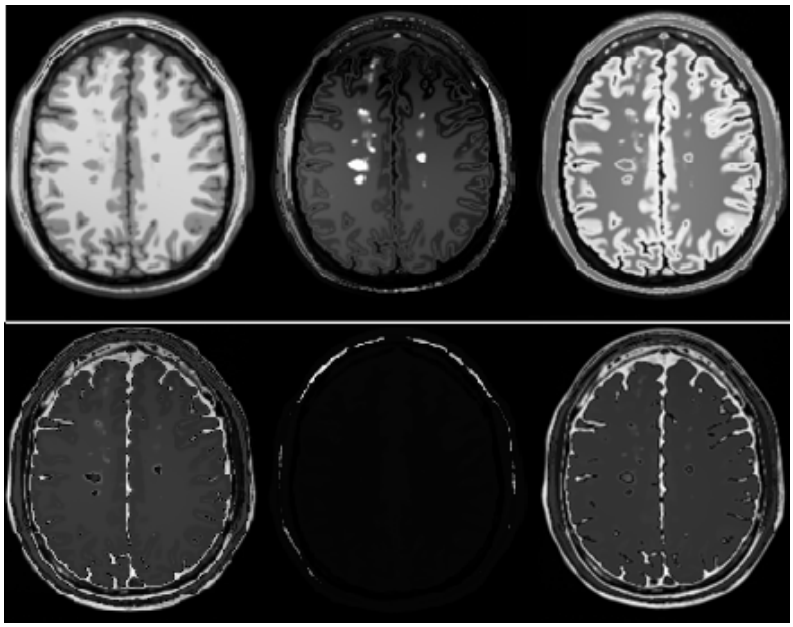
For example, Fig 4.11(a) shows sample slices, T1WI, T2WI, and PDI, representing MS lesions as input images. Multisignal wavelet analysis using *db4* wavelet is applied on the dataset. Reconstructed images from detail coefficients are shown in Fig. 4.11(b), in which source signals are decomposed into independent bases representing WM, CSF and GM from

left to right. Local features (here, MS lesions) are available from reconstructed image from approximation coefficient as given in Fig. 4.11(c). As a next step, FASTICA is applied on the expanded data set with number of bands extended to six, as explained in the next section.

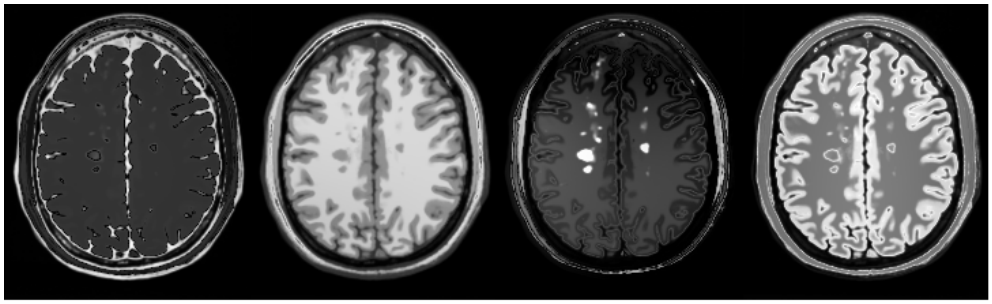
4.4.2 Feature extraction from expanded dataset

As we discussed in Chapter 3, band expansion with wavelet coefficients could unmix more brain tissues into six ICs as shown in Fig 4.12(a). Signals representing CSF in reconstructed images and input images generated two separate components yielding CSF for feature extraction. Some extrameningial tissues also got unmixed as a result of the new method. Relevant samples of CSF, GM, WM and lesions for feature selection are available from selected independent components given in Fig 4.12(b). ICs from conventional ICA (see Fig. 4.12 (c)) are included in this section for a comparative study with the generated ICs from MW-ICA.

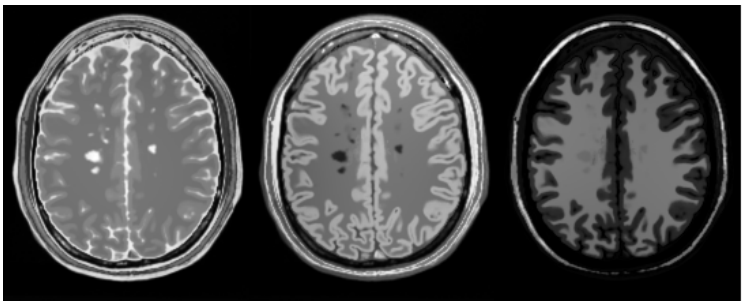
First of all, MW-ICA could unmix four brain tissues in separate ICs as shown in Fig. 4.12(b), which may provide better performance in classification. It is observed from conventional ICA results that features of CSF and WML are accumulated in 1st IC, whereas proposed method clustered CSF only. Comparing ICs representing WM (2nd IC in Fig. 4.12 (b) and 3rd IC in Fig. 4.12 (c)), presence of WML locations in WM is obvious in MW-ICA results. Compared to WML features from ICA results (1st IC in Fig.4.12 (c)), abnormality locations provided by 3rd IC in Fig. 4.12(b) is found to be more accurate. GM, represented by the last IC in Fig. 4.12(b) is good, but patches of abnormal tissues in it seem to be awkward compared to the results from ICA. A detailed analysis of real cases is required to ensure the effect of the algorithm in classification, which will be discussed in Chapter 5. The entire procedure of MW-ICA algorithm can be summarized in a few steps, as shown in Algorithm 4.3.



a. Independent Components from MW-ICA



b. Selected components from MW-ICA



c. Independent Components from ICA

Fig. 4.12 Independent Components from MW-ICA and ICA

Input: MRI Images I_1, I_2, I_3, I_4 etc. representing co-registered T1WI, T2WI, PDI, FLAIR etc.

Step1: Create p-dimensional multispectral image cube M = [I_1, I_2, I_3, I_4…I_p] from ' p' input images.

Step2: Reshape each image into its 1-D form and generate a p-dimensional multisignal **X**.

Step3: Apply multisignal wavelet analysis with level of decomposition to the greatest integer contained in ($\log_2(p)+1$) on **X** to decompose the signals into approximation coefficients and detail coefficients.

Step4: Apply multisignal wavelet reconstruction algorithm on detail coefficients to obtain reconstructed signal $\mathbf{X}^*$.

Step5: Expand **X** by appending $\mathbf{X}^*$ and form new input multisignal $\mathbf{X}_{new}$ of spectral dimension 'n' ≤ 2p.

Step6: Apply ICA on $\mathbf{X}_{new}$ to generate 'n' unmixed components. Reshape each c omponent to corresponding 2-D form.

Output: Enhanced independent components in 2-D form.

Algorithm 4.3 Proposed MW-ICA algorithm

4.5 Computational complexity and time analysis

Computational overhead is an important factor in dealing with the algorithms for clinical MRI analysis. From the literature, we studied that the conventional supervised classification approaches like neural networks, Bayesian classifiers, and maximum likelihood estimators are computationally intensive processes [24]. An optimal feature extraction procedure can help the classifiers to execute fast, but additional burden is to be counted on account of the feature extraction algorithms. Spectral clustering and wavelet analysis are the main concepts we introduced in the new methods. Following sections explains the main parameters affecting the computational complexity of the proposed methods.

4.5.1 Complexity of spectral clustering

In a study on number of operations required for computation of dissimilarity between two spectra with d bands using SAM, Paclík *et al.* (2003) [164] suggested that computational complexity is O(d). They noted that number of additions/subtractions is 2d FLOPS (FLoating point Operations per Seconds), and multiplications also can be performed in 2d Flops, with 1 division and 1 'arccos' computation as given by eq. (4.1).

In the proposed SC-ICA algorithm, number of clusters, C, has great impact on computational complexity. Decreasing the threshold will generate more clusters, which will increase the number of clusters also. Another parameter is the number of spectral signatures, N, included in the analysis. Considering all these, complexity of spectral clustering algorithm can be measured as (k_1xNxCxd), where k_1 is a constant.

4.5.2 Complexity of multisignal wavelet analysis

Computational cost for performing a wavelet transform for a filter of length L and a signal of length N=2^k is approximately measured as O(Nxd) [55], where d is the number of bands. In MRI analysis, size of the image and number of bands often found as fixed parameters. Wavelets in analysis often differ in length of the filter or number of vanishing moments, which linearly affects the execution time as shown in Table 4.3.

FASTICA computation time was measured in terms of number of samples and spectral bands, in the order of O(dxN). But, actual implementation using efficient techniques like code optimization and fixed point algorithms makes it very fast, apart from theoretical computations [144].

Table 4.3 shows time analysis results in seconds for the proposed feature extraction techniques. In this measurement we set d and N as fixed, and algorithm specific parameters like spectral threshold and wavelet type were

changed. Typical execution time required for a clinical abnormal case in PC with Pentium Dual CPU\2.0GHz\2GB RAM\Microsoft Windows 7\Matlab 7 configuration is available from Table 4.3.

Table 4.3 Time analysis

Methods	**Threshold α (clusters)**	**Time (sec.)**
SC_ICA	0.03 (833)	780
	0.06 (234)	222
	0.10 (81)	82
	0.15 (41)	40
	Wavelets	**Time (sec.)**
MICA	db2	1.4
	db4	2.2
	db8	4.9
	db16	11
MW-ICA	db2	2.7
	db4	3.9
	db8	7.7
	db16	20
ICA		0.7

For SC-ICA, complexity was measured on varying the threshold values (or the count of clusters). Observed results indicate the influence of number of clusters in execution time with a linear decrease on increasing the threshold (decreasing the number of clusters). For 833 clusters, time measured was 780, and 40 seconds were observed for 41 clusters, in effect showing a linear relationship between clusters and time for all threshold values.

For MICA, we estimated the time by varying the order of different daubechies wavelets. For db2, 1.4 seconds was measured, whereas db16 took 11 seconds for feature extraction. For db4 and db6 also, time variation was found to be linearly dependent on the number of vanishing moments. Compared to MICA, more time required for feature extraction using MW-ICA because of the additional overhead of ICA due to expanded data set. However, variation with different orders found to be linear, showing 2.7 seconds for db2, and 20 seconds for db16. In comparison with conventional ICA, it is observed that additional time required for feature extraction by wavelet based techniques varies from minimum 0.7 seconds to maximum 19.3 seconds approximately. SC-ICA is found to be very slow compared to all other methods.

4.6 Classification and segmentation

In this thesis work, effectiveness of the proposed algorithms in classification and segmentation is analyzed with widely used methods, FCM in the unsupervised category and SVM in the supervised category.

4.6.1 Support Vector Machines (SVM)

Support Vector Machines (SVM) is an efficient (non-linear) tool for supervised classification and regression. It is a linear discriminant function which was originally developed in statistical machine learning theory by Vapnik (1998) [28] as a linear binary classifier based on the class of hyper-planes,

$$(\mathbf{W}.\mathbf{x}) + \boldsymbol{b}=0,\ \mathbf{W} \in \boldsymbol{R}^{N},\ \boldsymbol{b} \in \boldsymbol{R} \qquad (4.6)$$

and decision functions

$$f(\mathbf{x}) = \text{sign}((\mathbf{W}.\mathbf{x}) + \boldsymbol{b}) \qquad (4.7)$$

where '**W**' is a weight vector and '***b***' is the threshold or bias. SVM searches for an optimal hyper-plane having maximal margin of separation between two classes for a particular training dataset. Therefore, the classification task is only a function of the *support vectors*, the training data that lie on the margin.

In dual form [28] the problem reduces to

$$\text{Maximize} \sum_{i-1}^{n} \alpha_i \; - \; \frac{1}{2} \sum_{i=1}^{n} \alpha_i \alpha_j y_i y_j \; K(x_i, xj) \qquad (4.8)$$

Subject to $\alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i y_i = 0$

and solve for α_i. $K(x_i, x_j)$ can be a linear or non-linear kernel.

For a particular training data set, there may be several hyper-planes corresponding to maximum margin separation between two classes as shown in Fig. 4.13. SVM searches for an optimal hyper-plane, satisfying eq. (4.6) with the maximal margin of separation between the two classes. Optimal hyper-plane is orthogonal to the shortest line connecting the convex hulls of the two classes and intersecting it half way. Support vectors are elements of the training set that lie on the boundary hyper-planes of the two classes as shown in Fig.4.13 (annotated with circles). The main attractive feature of SVM is that classes which are nonlinearly separable in the original space can be linearly separated in the higher dimensional feature space F via a nonlinear map φ as shown in Fig 4.14 and denoted by

$$\varphi \rightarrow R^N \rightarrow F \qquad (4.9)$$

This only requires the evaluation of dot products

$$k(\mathrm{x},\mathrm{y}) = (\varphi(\mathrm{x}).\varphi(\mathrm{y})) \qquad (4.10)$$

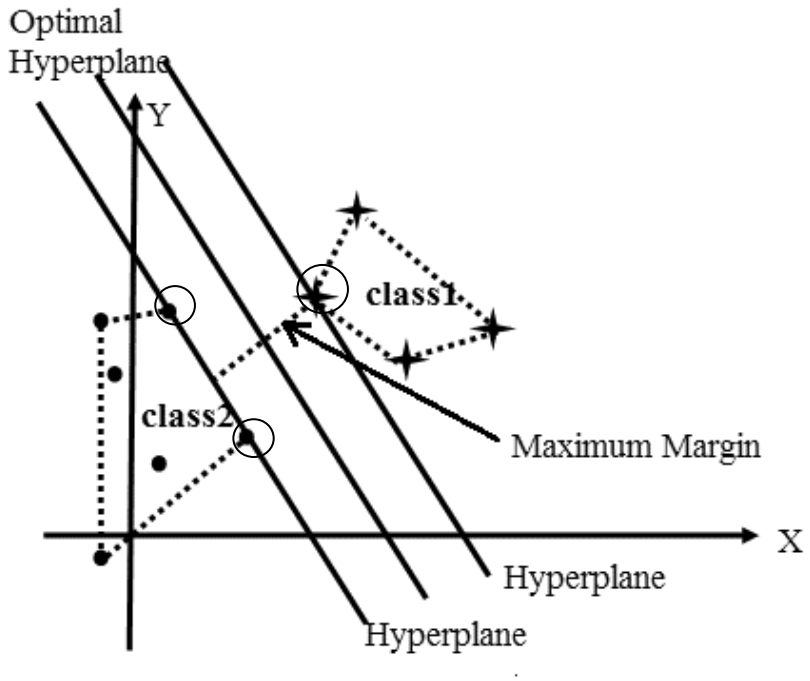


Fig. 4.13 Support vectors (circled points) and SVM classification

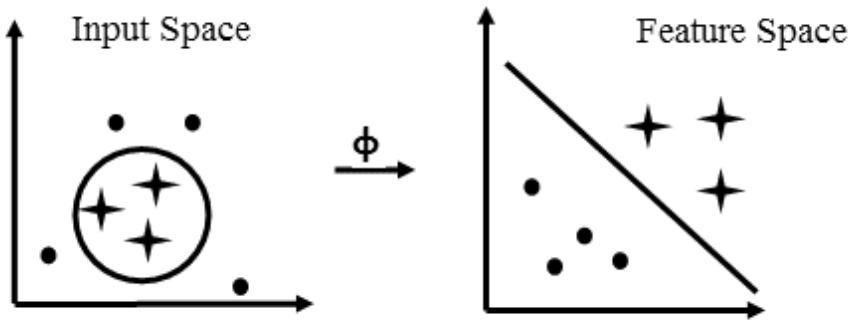


Fig. 4.14 The idea of nonlinear SVM: map the training data nonlinearly into a higher-dimensional feature space F via φ.

A great number of kernels exist in literature like RBF, polynomial, quadratic etc. [26]. The concept of kernels in SVM makes it capable of solving complex nonlinear classification problems.

Important characteristics of SVM are its ability to solve classification problems by means of convex Quadratic Programming (QP) and also the sparseness resulting from this QP problem. The learning is based on the principle of structural risk minimization. Instead of minimizing an objective function based on the training samples (such as mean square error), the SVM attempts to minimize the bound on the generalization error (i.e., the error made by the learning machine on the test data not used during training). As a result, an SVM tends to perform well when applied to data outside the training set. SVM achieves this advantage by focusing on the borderline training examples, *support vectors* [68].

It has been proved in latest studies that SVM is a better option for MRI analysis [29, 30] compared to other methods in supervised classification. SVM with Radial Basis Function (RBF) non-linear kernel is used in this thesis for supervised brain tissues analysis.

4.6.2 Fuzzy C-Means Clustering (FCM)

It is a data clustering technique introduced by Bezdec in 1981[120], where each data point belongs to a cluster to some degree that is specified by a fuzzy membership grade [117]. Let $\mathbf{X}=(x_1, x_2,.,x_N)$ denotes an image with N pixels to be partitioned into 'c' clusters, where x_i represents multispectral (features) data, and 'c' is the number of clusters with $2 \leq c <$ n. The standard FCM objective function for partitioning a dataset $\{x_k\}_{k=1}^{N}$ into 'c' clusters is given by,

$$\mathrm{J_m} = \sum_{i=1}^{c}\sum_{k=1}^{N} U_{ik}^{m} \left\| x_k - v_i \right\|^2 \tag{4.11}$$

where $\{v_i\}_{i=1}^{c}$ are the centres of the clusters and

the array $\{U_{ik}\}, i = 1..c, k = 1,..N$ is a partition matrix such that

$$U \in \left\{ U_{ik} \in [0,1] \middle| \sum_{i=1}^{c} U_{ik} = 1, \forall k \ \ \& \ 0 < \sum_{k=1}^{N} U_{ik} < \mathrm{N}, \ \forall i \right\}$$

The parameter '*m*' is a weighting exponent on each fuzzy membership, and it determines fuzziness amount of the resulting classification. Gray level values are the most commonly used feature in image processing. So FCM objective function $\mathrm{J_m}$ is minimized when high membership values are assigned to pixels whose intensities are close to the centroid of its particular class, and low membership values are assigned when the point is far from the centroid [109]. In the FCM algorithm, the probability that a pixel belongs to a specific cluster depends only on the distance between the pixel and each individual cluster center in the feature domain. Algorithm starts with an initial guess for each cluster centre, and it converges to a solution for v_i representing the local minimum or a saddle point of the objective function, $\mathrm{J_m}$. A detailed explanation of FCM theorem and algorithm is available in [104].

4.7 Summary

As an attempt to resolve the limitations of ICA in multispectral brain MRI analysis, we proposed a few extensions to ICA, based on wavelet analysis and spectral clustering. In this chapter we discussed three new methods to improve the performance of brain tissue classification from brain MRI; a new Spectral Clustering ICA (SC-ICA), a new Multisignal Wavelet ICA (MW-

ICA) and modified Multiresolution ICA (MICA). A detailed explanation of each method, including the relevant theory, algorithm and examples are provided in this chapter. Computational complexity on varying different parameters is also discussed for the proposed methods. To justify the application of the new methods in real environment, a comparative study of execution time in seconds for a clinical dataset is added to the end of the chapter. A detailed performance analysis with a large size database is required to evaluate and confirm the efficiency and potential of these methods in supervised and unsupervised brain tissue classification. In the next chapter, we revisit all of these algorithms as new methods for efficient classification using FCM and SVM.

Chapter -5

BRAIN MRI ANALYSIS WITH PROPOSED ICA EXTENSIONS

5.1 Introduction

Chapter 4 proposed three ICA extensions based on spectral distance and wavelet transforms for improved feature extraction. This chapter analyzes the effectiveness of these methods in brain tissue classification using supervised and unsupervised approaches. Support Vector Machines (SVM) was selected for supervised segmentation, because of its generalization capability and efficiency with a few features and small training set. Fuzzy C-Means clustering (FCM) was considered for unsupervised segmentation, since it is robust to ambiguity and efficient in retaining much more information than hard segmentation methods. A detailed explanation on the database used in the evaluation of the algorithms, experiments conducted and performance evaluation is included in Section 5.2. Section 5.3 explains the preparation of input data for image analysis. Section 5.4 discusses the measurements for performance evaluation. Quantitative and qualitative analysis of results and discussions are included in Section 5.5.

5.2 Database and experimental setup

5.2.1 Input data

Total 120x3 axial slices from T1-Weighted Images (T1WI), T2-Weighted Images (T2WI), and Proton Density Images (PDI) with parameter settings 1-mm slice thickness, intensity non-uniformity 0% and noise level 0% from

Brainweb[20] database were included in the synthetic image analysis. Out of 181x3 normal slices, 60 slices were considered from each sequence to form the healthy brain data, and 60 out of 181 abnormal slices giving information on multiple sclerosis (MS) lesions were selected to give abnormal multispectral images. The groundtruth images available from Brainweb database were utilized for performance comparison.

The normal and abnormal brain MR images from T1W sequence, T2W sequence and FLAIR sequence from total 382 patients with specifications as described in Chapter 1 were utilized for clinical evaluation. 168 cases were identified as normal, 110 cases were diagnosed with White Matter Lesions (WML), and 104 were identified with other abnormalities like tumor, lesions, edema etc. No gold standard exists as groundtruth for comparison of the results. The groundtruth tissues in clinical cases were collected from manually segmented and labeled images, after removal of extrameningial tissues, under the supervision of an experienced radiologist.

5.2.2 Experiments

For the supervised classification, the entire database was divided into two sets: training data and test data, as given in Table 5.1. Synthetic database were obtained as 181 slices of a normal or abnormal cases with same parameter settings. 60 slices selected were distributed as 40 for training and 20 for testing. In the case of healthy clinical cases, 120 cases were considered for training and 48 were selected for testing. In the clinical abnormal category, 150 cases were included in the training set and 64 were counted for testing. After image registration and data preparation, four sets of experiments were performed on each database. First, feature extraction using ICA and proposed techniques (as described in Chapter 4) was performed. The second experiment was the segmentation of test data using FCM. Training, model selection and classification were included in the third set;

fourth set was the qualitative analysis of the classified tissues by FCM and SVM. In the third set of experiments, feature vectors were collected with the help of a 3x3 window under the guidance of an experienced radiologist. Total number of feature vectors collected for testing and training are given in Table 5.1.

Table 5.1 Validation and Test plan

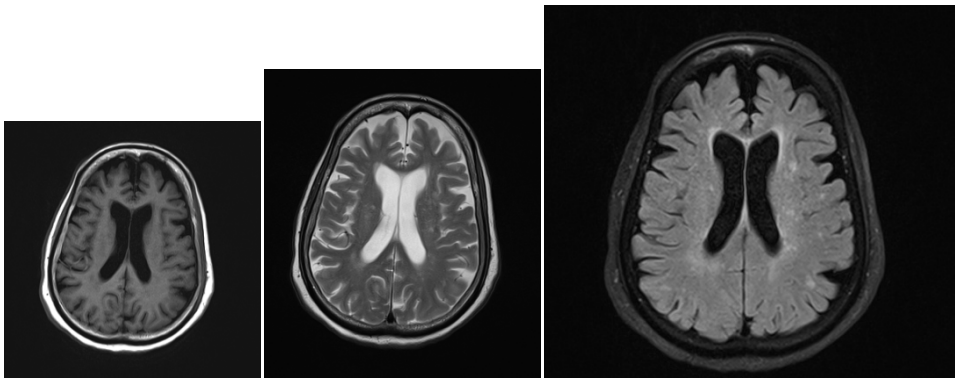
Category	Training data count		Cross validation	Testing data count		Total feature vectors
	patients	Slices		patients	Slices	
Synthetic normal	1	40	External 10-fold	1	20	1800
Synthetic abnormal	1	40		1	20	1800
Clinical normal	120	360		48	144	7560
Clinical abnormal	150	450		64	192	12840

5.3 Data preparation for multispectral analysis

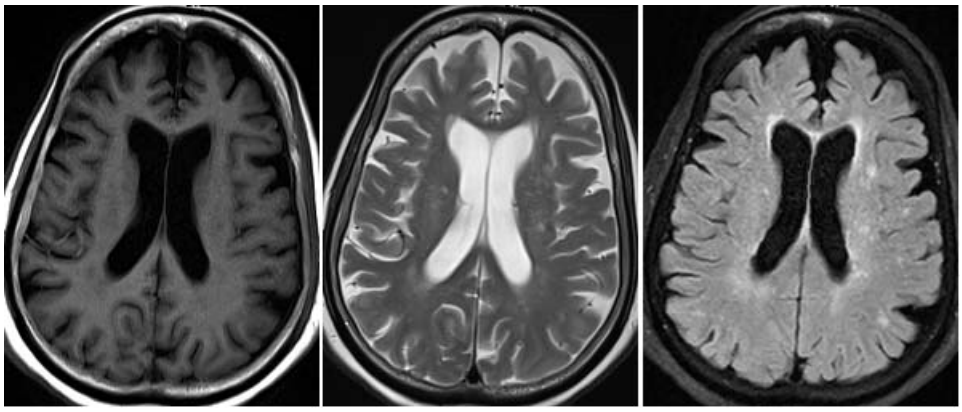
As we discussed earlier in Chapter 1, MR sequences must be registered to the accuracy of the pixels before selecting them for a multispectral analysis. It is also noted that analysis is usually performed on the spectral signatures corresponding to every pixels. Input MRI images collected from 3T machines were of very good quality. Pre-processing steps such as contrast enhancement, noise removal or any other signal modification on these images were observed as sometimes eliminating small, but significant pathological information. So they are avoided from the image analysis in this thesis work.

5.3.1 Image registration

Images in synthetic database were obtained in the form of co-registered. i.e the corresponding slices in T1-weighted, T2-weighted and PD sequences were obtained with the same direction and same size, and pixel vectors were found to be collinear.



(a) Original Clinical Sequences, in the order of T1WI, T2WI and FLAIR image from left to right



(b) Co-registered clinical sequences

Fig. 5.1 Image registration of clinical images

But clinical sequences collected were of different size and direction. Image resizing and rotation was enough for the majority of the clinical sequences in this work. In certain cases, we used control points based image registration using MATLAB functions also. Fig. 5.1 illustrates the registration in clinical sequences. The original clinical sequences, shown in Fig. 5.1(a), are of different resolution and direction. Images in Fig 5.1 (b) were generated as a result of rotation, resizing and control points based manual registration using Matlab functions. After registration, size of the registered images was fixed to 209x276 pixels and 227x260 pixels for normal and abnormal cases respectively.

5.3.2 Spectral signature collection and feature extraction

This thesis work is based on multivariate analysis using ICA. So we prepared the data for an analysis with reduced complexity, by reshaping each 2-D image into 1-D array. We got three 1-D arrays from three bands, which were combined to form a matrix with columns representing each pixel vector, as given in Fig. 5.2. Intensity values of selected window are also plotted in Fig. 5.2

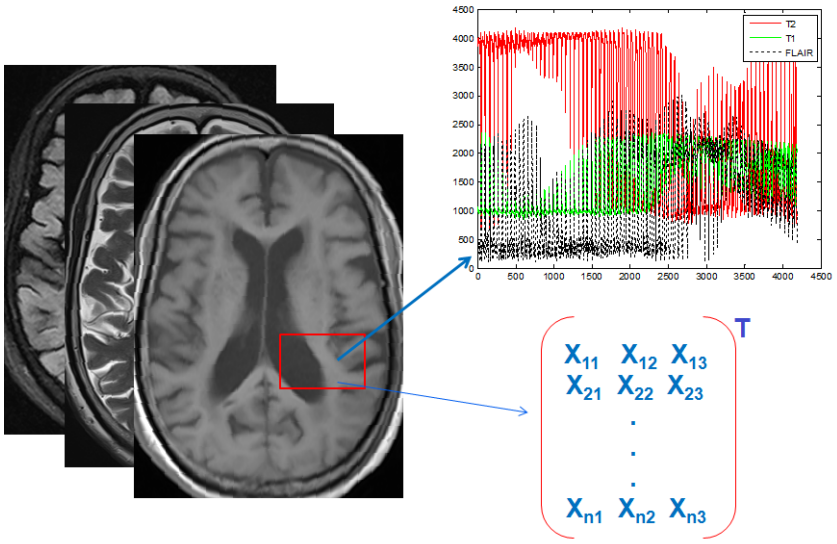


Fig. 5.2 Spectral signature sets

Spectral signature of n^{th} pixel can be given by pixel vector $[X_{n1}\ X_{n2}\ X_{n3}]^T$, for n=1,2,3…. Each multispectral cube discussed in this work is described as a collection of these spectral signatures. Multivariate analysis based on ICA and proposed ICA extensions were performed on this collection, as we discussed in Chapter 4. Fig. 5.3 shows ICs from ICA representing WM and WML; annotated portions shows 3x3 Feature Vectors (FV) selected for supervised analysis. For unsupervised classifications, no feature selection was required; it was performed directly on the generated ICs.

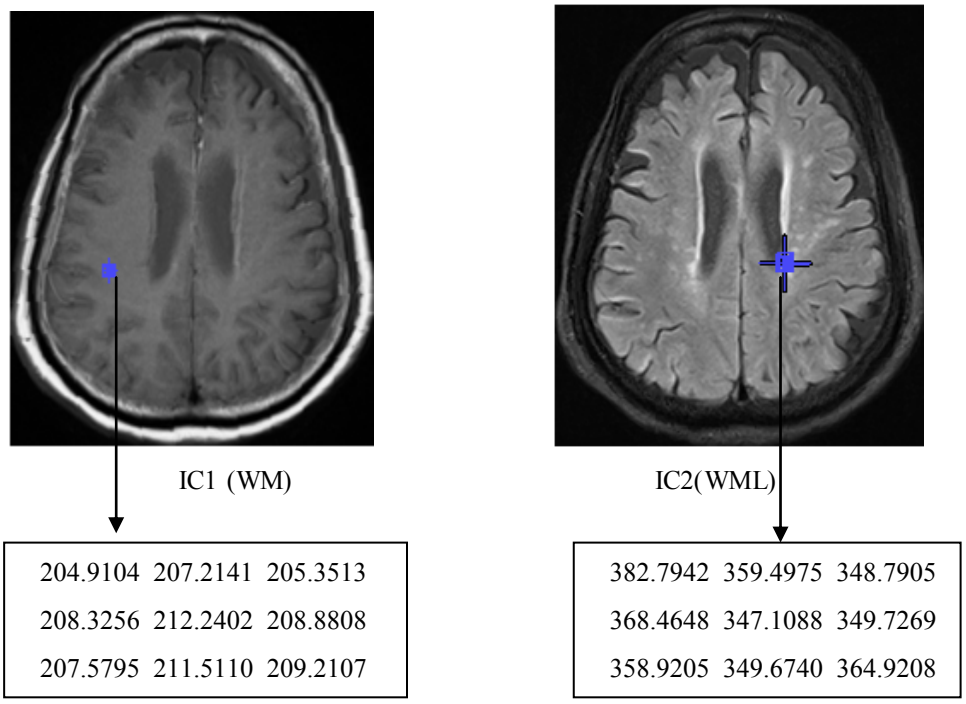


Fig 5.3 Feature vector selection of WM and WML from ICA results

5.3.3 Bias estimation for SVM classification

A bias estimation on varying number of feature vectors was performed with SC-ICA based SVM to select the best validation technique. Cross Validation (CV) [77] techniques, Internal CV Leave One Out (ICV LOO), external 50% Holdout (EHoldout) and External 10-fold CV (ECV 10) were considered for the bias estimation. We followed the selection bias assessment in Ambroise and McLachlan, (2002) [165] to conduct this study. They suggested that external 10-fold CV, in which the information from test data is not used for the development of the models, can be used for reasonable assessment of accuracy of predictions. In gene array classification, there may be more than 10000 features to define a gene. Therefore, they have done the experiments with changing number of features.

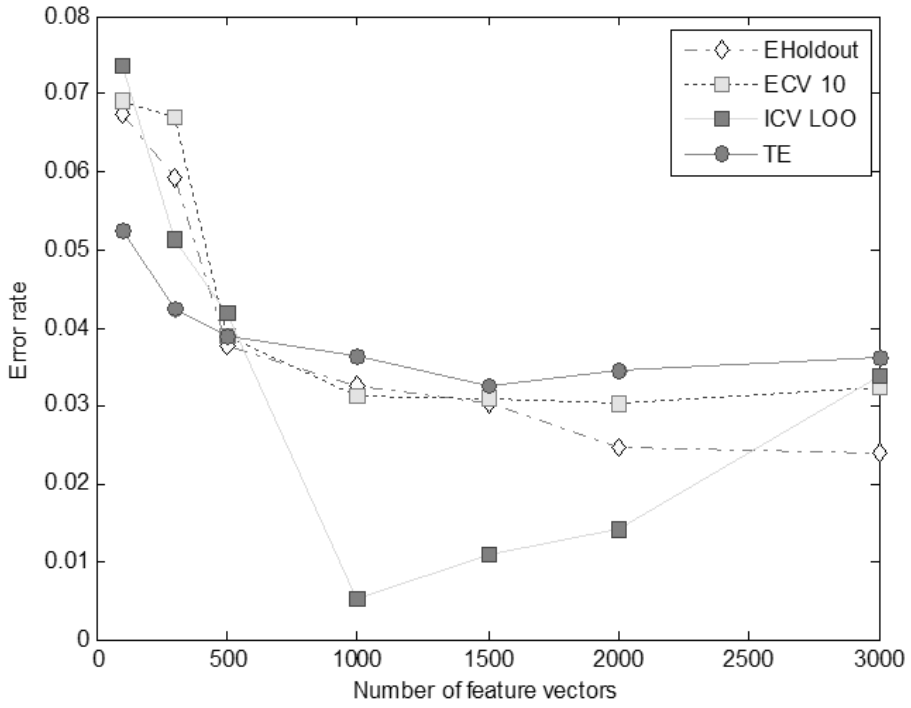


Fig. 5.4 Bias estimation for internal and external cross validations

However, in the case of MRI analysis, we conducted the experiment by varying the number of feature vectors. Error estimates were first observed for 3000 feature vectors in 100 random splits. Same procedures were repeated for different number of feature vectors 2500, 2000, etc., and average error rates were estimated as plotted in Fig. 5.4. Results obtained for unbiased prediction error (TE) were also added to study the error variation. A detailed explanation of selection bias and error rate estimation is available in [77, 165]. The models obtained on each fold of the ICV LOO were based on the same set of feature descriptors, whereas corresponding models for the external CV involved different feature descriptors. In Fig. 5.4, ICV LOO results were found to be more biased compared to EHoldout, whereas ECV 10 and TE keep almost similar behavior. So ECV10 validation was selected for supervised analysis throughout this thesis work.

5.4 Performance evaluation

Performance evaluation of the proposed classification was carried out with quantitative and qualitative analysis. Both approaches have strengths and weaknesses; a particular strength of quantitative research is that statistical analysis allows for generalization to entire population. However, qualitative results provide depth and richness of information, which is not possible with the quantitative data. Qualitative imaging and image analysis are normally used in a clinical environment, which is performed often by visual assessment of MR sequences. The signal intensity may vary between patients and measurement sessions [166]. Sometimes this may arise inter operator variations due to lack of expertise in the field.

Tanimoto Index, sensitivity, specificity, accuracy, False Positive Rate (FPR), False Negative Rate (FNR) and Error rate were the measures used in quantitative analysis in this work. Tanimoto Index, the most commonly used criterion in medical imaging [5, 29] is used to measure the similarity of the classified tissues with the groundtruth values, and it is given by,

$$\mathrm{TI} = \frac{|A \cap B|}{|A \cup B|} \tag{5.1}$$

where A is the dataset representing the classified brain tissues like CSF, GM, WM, Tumor, etc., and B is the corresponding ground truth dataset.

Classifier performance evaluation in this work is conducted with widely used statistical measures, sensitivity, specificity, accuracy etc. [77] given by,

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{5.2}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{5.3}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.4}$$

Error rate = 1- Accuracy (5.5)

FPR = 1 – Specificity (5.6)

FNR = 1 – Sensitivity, (5.7)

Where True Positive (*TP*) is defined as the number of correctly identified positive pixels; True Negative (*TN*) is defined as correctly identified negative pixels. For example, in a diagnostic test evaluation focusing on the presence of abnormal tissues, tumor samples are considered in the positive category and normal tissues will be in the negative category. False Positive (*FP*) represents the count of normal tissues incorrectly identified as tumor, and False Negative (*FN*) gives the count of abnormal samples incorrectly identified as normal tissues. Higher values of sensitivity, the proportion of correctly classified positives, indicate the good performance of the method in predicting positives. Specificity measures how well the system can predict the negatives. Accuracy measures the overall correctness of the classifier in predicting both positives and negatives.

We used an additional method, Bland-Altman analysis [138], to evaluate the agreement between measures from two methods. In the Bland-Altman plots, difference between two measurements (the bias) is plotted against the average of those measurements. In this work, estimated lesion volumes from classification by ICA and proposed methods were quantitatively compared with lesion volumes in the groundtruth.

The best classification models selected from validation procedure in supervised learning were used to classify the brain tissues. These classified images were included in the qualitative analysis of supervised classification. Segmented images from FCM based on ICA or ICA extensions were first labeled by an experienced radiologist, and visually compared with groundtruth images.

5.5 Result analysis and discussion

Analysis of experimental results is divided into two sections: Synthetic image analysis and Clinical image analysis. Mathworks Matlab 7.0 (R2009a) implementation of the algorithms on a PC with Pentium Dual CPU of 2.1GHz and 4GB RAM running Microsoft Windows 7 was executed for the complete system evaluation. Pattern recognition Toolbox in Matlab is used for non-linear (RBF kernel) SVM training and classification, with default parameter settings, and Sequential Minimal Optimization (SMO) as optimization method. One-against-all SVM strategy was adopted to solve the classification problem. FCM method provided in Fuzzy Logic Toolbox was applied on generated ICs to do automatic unsupervised segmentation.

5.5.1 Synthetic image analysis

For the first set of experiments, feature extraction, were performed on synthetic normal and abnormal training data, by applying proposed and conventional methods in the order of SC-ICA, MICA, MW-ICA and ICA. SC-ICA was evaluated with different threshold values, 0.1, 0.06, 0.03 etc. Daubechies wavelets of different orders were considered in multisignal wavelet analysis. CSF, GM, WM, and WML were the main brain substances to be classified from synthetic database. Feature vectors from training and test data were generated from normal and abnormal synthetic database as described in section 5.3, and utilized in experiments 3, and 4 for a detailed performance evaluation.

5.5.1.1 Unsupervised classification, FCM

20 test cases as given in Table 5.1 were considered to evaluate the performance of new methods with unsupervised segmentation using FCM. Not much variations observed in the results from test data. Average of Tanimoto Index (TI) values from test datasets are summarized in Table 5.2.

Results from FCM segmentation based on SC-ICA (SC-ICA+FCM) for different thresholds, FCM based on MICA (MICA+FCM) and MW-ICA (MW-ICA+FCM) for wavelet *db12*, and ICA based FCM are included in Table 5.2. SC-ICA based segmentation for thresholds 0.1, 0.06, and 0.03 recommended 0.06 as the best threshold value yielding highest TI values, 0.78/0.72/0.93, for normal CSF/GM/WM, and 0.81/0.85/0.94/0.89 for abnormal CSF/GM/WM/WML.

Table 5.2 Tanimoto Index from synthetic image analysis

		Normal				Abnormal			
Methods	α	CSF	GM	WM		CSF	GM	WM	WML
SC-ICA+ FCM	.1	0.72	0.69	0.92		0.81	0.81	0.93	0.84
	.06	0.78	0.72	0.93		0.81	0.85	0.94	0.89
	.03	0.76	0.67	0.91		0.76	0.79	0.97	0.34
MICA + FCM		0.85	0.80	0.65		0.86	0.79	0.68	0.94
MW-ICA+FCM		**0.84**	**0.79**	**0.94**		**0.84**	**0.85**	**0.95**	**0.91**
ICA + FCM		0.71	0.73	0.85		0.76	0.84	0.92	0.87

WML classification for threshold 0.03 yielded a low TI value, 0.34, but at the same time providing best value, 0.97, for WM classification. It may occur because of false positives present in the classified result due to over clustering. We will discuss those details in Chapter 7.

MICA+FCM could perform better for all tissue classes, except for WM. In the feature extraction phase, we noted that IC corresponding to WM was of low quality. It is well reflected in the observed low TI value also. However, TI for WML classification is found to be the best, 0.94, among all the results in Table 5.2. MW-ICA+FCM could give reasonably high TI values for all brain tissues. Overall performance of MW-ICA+FCM recommends it as the best method for segmentation. Comparison with conventional ICA based results, gives the improvement provided by each method. All except MICA+FCM showed better results for all brain tissues; MICA based results

failed to reach the performance of ICA based classification in the case of WM.

The effectiveness of these methods in WML classification was further evaluated with Bland-Altman plots as shown in Fig. 5.5 - Fig 5.8. The differences between lesion volumes (in mL) from category A and category B were plotted against the averages of both measures to generate the plots. WML volumes calculated from the classified results from the proposed methods and conventional ICA (category B), are compared against the lesion volumes provided by category A, the groundtruth images in Brainweb database. 20 synthetic abnormal multispectral sets were considered in this evaluation. SC-ICA was performed with threshold value 0.06, and *db12* was selected for wavelet analysis. The mean difference (bias or *d*), Standard Deviation (SD) of the differences, and 95% limits of agreement (± 1.96 SD) were used to analyze the plots. Considering the classification results from each method in category B, the following results were observed.

- Using ICA based classification, *d* was 0.24mL with 95% limits of agreement, -0.11mL to 0.58mL (Fig. 5.5), whereas the points were distributed closer to zero for SC-ICA based segmentation. Bias was observed as 0.04 mL with 95% limits of agreement, *[-0.02mL, 0.11mL]* (Fig. 5.6). Number of outliers is reduced to one, and confident interval also got shortened.
- However, MICA+FCM could not provide such a performance in lesion analysis. We observed a higher *d* value 0.06 mL, and a wider confidence limit, *[-0.06mL, 0.18mL]* (Fig. 5.7), compared to SC-ICA+FCM.

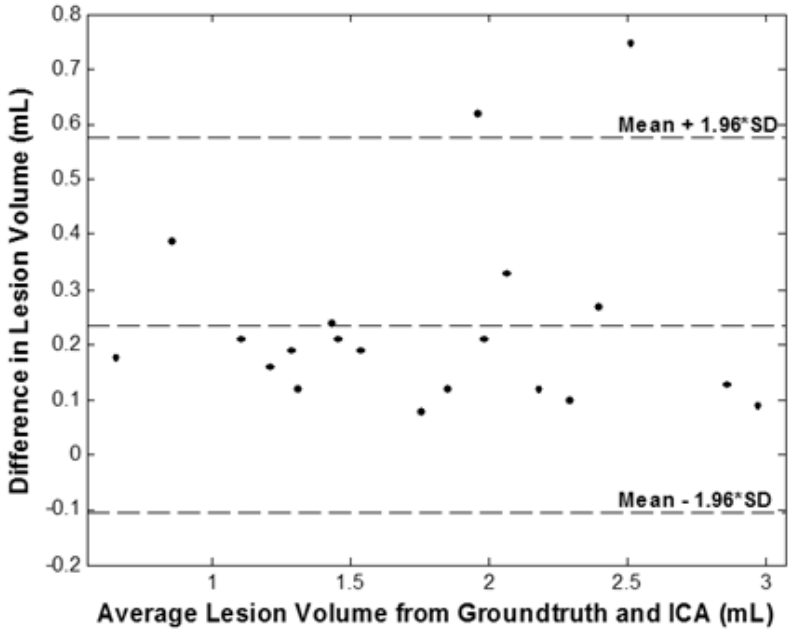


Fig 5.5 Bland-Altman plots based on lesion volumes between groundtruth and ICA based FCM. The observed bias, *d = 0.24mL*, and 95% confidence limit = *[--0.11mL, 0. 0.58mL]*

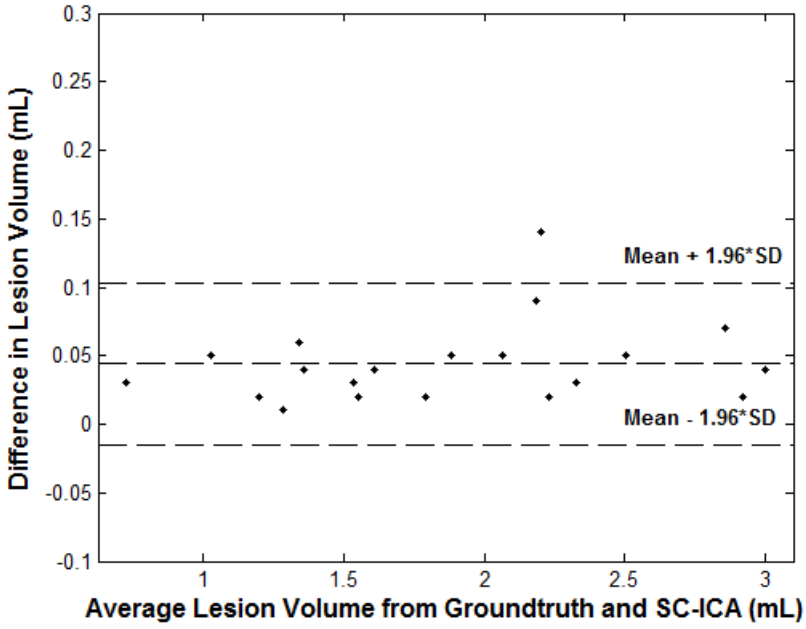


Fig 5.6 Bland-Altman plots based on lesion volumes between groundtruth and SC-ICA based FCM. The observed bias, *d = 0.04mL*, and 95% confidence limit = *[-0.02mL, 0.11mL]*

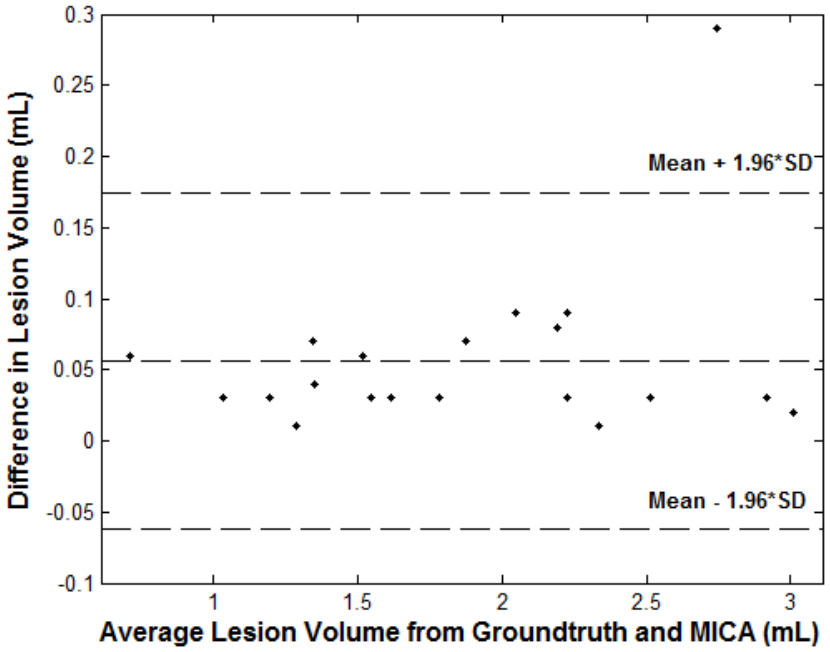


Fig. 5.7 Bland-Altman plots based on lesion volumes between groundtruth and MICA based FCM. The observed bias, *d = 0.06mL*, and 95% confidence limit = *[-0.06mL, 0.18mL]*

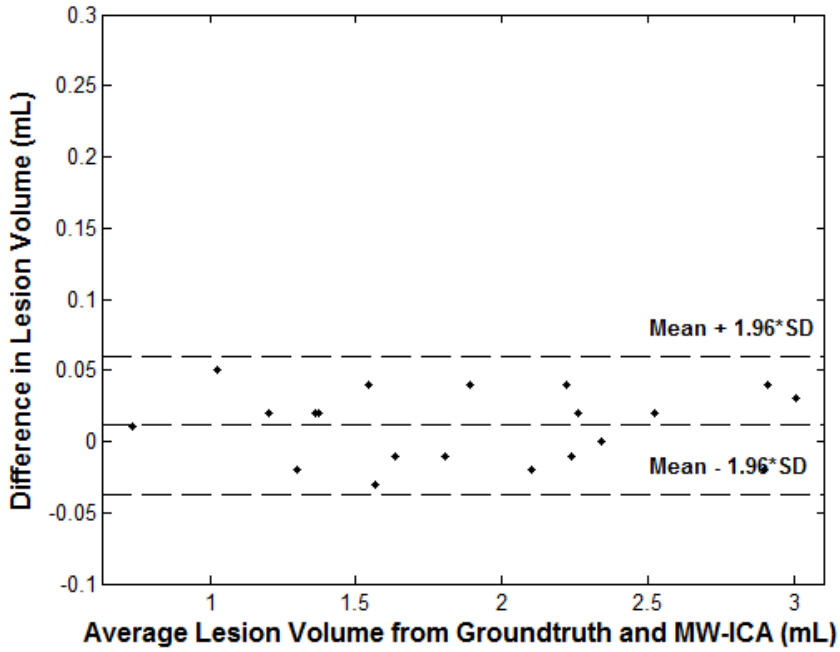
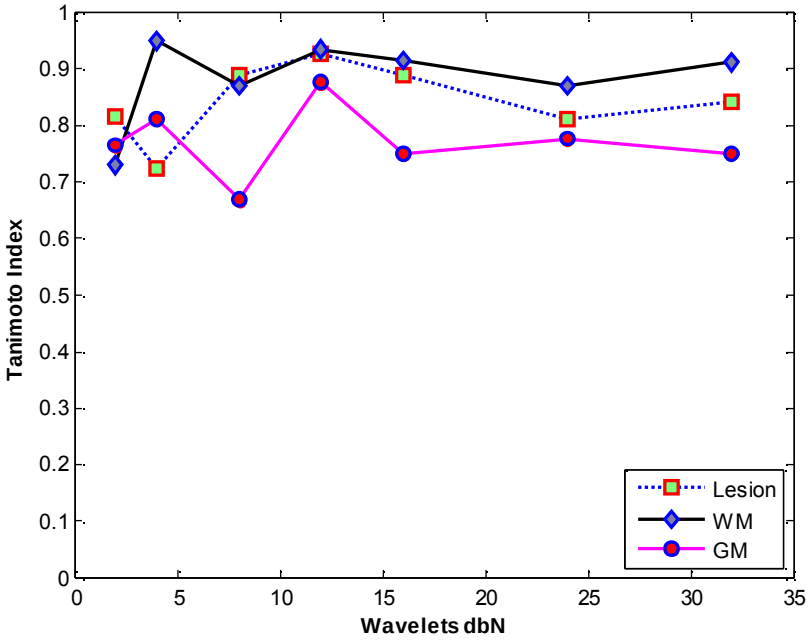


Fig. 5.8 Bland-Altman plots based on lesion volumes between groundtruth and MW-ICA based FCM. The observed bias, *d = 0.02mL*, and 95% confidence limit = *[-0.03mL, 0.06mL]*

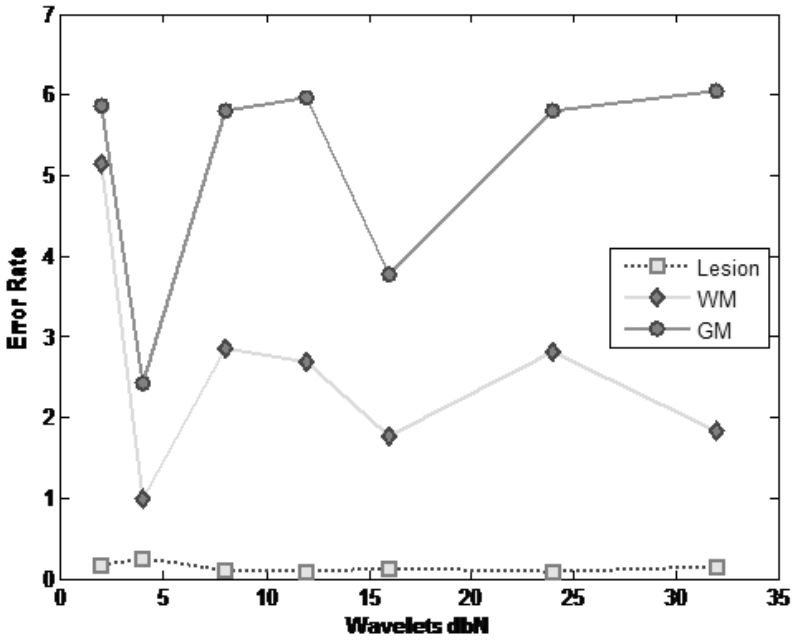
- The best performance was provided by MW-ICA based FCM (Fig. 5.8), with least bias, *d = 0.02mL,* and narrow agreement limit, *[-0.03mL, 0.06mL]*; i.e. bias was contained well within the two standard deviations, with no outliers observed outside the limits of agreement.

To confirm the efficiency and consistency of the new wavelet based ICA extensions, MW-ICA algorithm was evaluated for different wavelets, *db2, db4, db8, db12, db16, db24,* and *db32* for 15 synthetic cases in the test data. TI values and error rate estimated in percentage for GM, WM and Lesion (WML) were plotted as shown in Figure 5.9 (a) and 5.9 (b) respectively. X-axis shows Daubechies wavelets of different orders, '*N*' [163]. From Figure 5.9 (a), it is observed that TI value is low for db2 for the three classes considered. Classified lesions are found to be better on varying the order from 4 to 12, reached the maximum, above 0.92, at db12, and then showed a decrease for db16, db24 and db32. Db4 is selected as best, 0.95, for WM classification, whereas showing the next best value at *db12*, around 0. 92. For GM extraction, *db12* yielded the best results, around 0.87, and *db4* was observed as the next best choice with TI value around 0.82. All these results supported wavelets with low to medium regularities for best TI calculation.

TI values alone cannot decide the optimal parameters, since these values are affected by false positives in calculation. Therefore, we considered another measure, error rate (1-Accuracy) also to check the variation for different number of vanishing moments, as shown in Fig. 5.9(b). Db4 was observed with least error rate for GM and WM classification; GM classification showed an error rate of around 2.5%, whereas WM showed less than 1%. For lesions, error rate was found to be very less, nearing zero, for almost all *dbN* wavelets, *N= 2, 4, 12*…., with least value observed for db12. From all these observations, we considered db12 and db4 as candidates for synthetic image analysis.



(a). Tanimoto Index variations for different orders of daubechies wavelets



(b). Error rate variations for different orders of daubechies wavelets

Fig.5.9. Performance measure variations on varying order of daubechies wavelets

FCM results were found to be significantly affected by the uncertainties in the input images and membership calculations. Therefore, a more efficient classification method, especially a supervised learning and classification can be a good choice for better analysis of brain tissue classification.

5.5.1.2 Supervised classification, SVM

As given in Table 5.1, we considered total 1800 feature vectors representing different brain tissues, CSF, GM, WM and WML for supervised analysis. 1200 feature vectors were considered in the validation and model selection. Classification results using 600 feature vectors from 20 abnormal test sets are given in Table 5.3. Sensitivity, specificity and accuracy measured for GM, WM and WML are summarized in the results. SVM based on SC-ICA for different thresholds 0.1, 0.06 and 0.03, MICA and MW-ICA for *db12* wavelet, and ICA were considered in the evaluation procedure. The proposed methods were found to be yielding better results compared to ICA+SVM. It is observed that overall performance of MW-ICA+SVM is optimal among all these methods described in the analysis.

Table 5.3 Performance analysis of synthetic brain tissue classification

		GM			WM			WML		
Methods	α	Spec.	Sens.	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.	Acc.
SC-ICA+ SVM	.1	97.06	82.54	92.02	96.18	92.11	94.64	99.98	83.79	99.93
	.06	97.57	84.56	94.91	97.48	93.87	94.84	99.94	91.84	99.71
	.03	94.56	79.04	88.49	92.33	98.07	95.63	89.92	97.35	91.31
MICA+ SVM		89.98	79.87	88.64	98.41	67.75	90.21	99.91	94.92	99.82
MW-ICA+SVM		**97.03**	**84.85**	**93.94**	**96.92**	**97.84**	**97.04**	**99.94**	**94.23**	**99.72**
ICA+ SVM		94.87	84.28	91.92	98.0	92.53	97.91	99.91	85.88	99.61

Spec. = Specificity, Sens. = Sensitivity, Acc. = Accuracy

From Table 5.3 following highlight points are observed;

- In GM classification, SC-ICA_.06+SVM was found to be giving best accuracy, 94.91, maintaining good values for specificity and sensitivity. MW-ICA+SVM also performed almost same as SC-ICA_.06+SVM in GM classification. But results of MICA+SVM were found to be not satisfactory, because of the global suppression in the algorithm.
- MW-ICA+SVM exceeded all other methods in WM classification with highest sensitivity/accuracy values, 97.84/97.04. Sensitivity value provided by SC-ICA+SVM for threshold 0.03 was found to be the highest, but over-clustering effects reduced the specificity and overall accuracy.
- SC-ICA_.03+SVM showed a similar behavior in classification of WML also. True positives were best predicted with sensitivity value, 97.35, but specificity lowered to 89.92. MICA+SVM provided the best performance in WML analysis with very good specificity/sensitivity/accuracy measures, 99.91/94.92/99.82. MW-ICA+SVM also presented a similar performance in WML classification.

Sensitivity and specificity are good measures in assessing the performance of the classification system in identifying true positives and true negatives. But segmented tissues may show large deviation from these results due to vagueness and ambiguity in input data. Then, potential of the methods in segmentation can be better described by Tanimoto Index. The classified brain tissues from synthetic test data using SVM models were compared with the groundtruth tissues in Brainweb, and average of estimated TI values is tabulated in Table 5.4. Both normal and abnormal datasets were considered

to study the variation in classifying more tissue classes from less number of images.

Table 5.4 Tanimoto Index values from classified brain tissues

		Normal				Abnormal			
Methods	α	CSF	GM	WM		CSF	GM	WM	WML
SC-ICA+ SVM	.1	0.71	0.69	0.92		0.78	0.83	0.94	0.85
	.06	**0.77**	**0.78**	**0.94**		**0.79**	**0.85**	**0.94**	**0.82**
	.03	0.74	0.71	0.91		0.77	0.76	0.97	0.42
MICA + SVM		0.85	0.74	0.67		0.79	0.81	0.71	0.94
MW-ICA+SVM		**0.86**	**0.72**	**0.95**		**0.81**	**0.79**	**0.88**	**0.96**
ICA + SVM		0.73	0.74	0.84		0.72	0.83	0.86	0.84

From Table 5.4, proposed classification methods were found to be yielding better performance than ICA based classification. SVM based on SC-ICA_.06 and MW-ICA were selected as optimal methods to segment the brain tissues. Considering SC-ICA based classification, threshold value 0.06 was found to be giving best TI values for CSF/GM/WM, 0.77/0.78/0.94 for normal case. But abnormal case analysis showed that threshold value 0.1 also can be selected as a candidate to provide high performance classification. Classified WM tissues were found to be best for threshold 0.03, but over clustering adversely affected the classification of other tissues. However, introduction of multisignal wavelet analysis to ICA improved the TI values of small WML classification from 0.84 by ICA based SVM to 0.94 by MICA based SVM, and 0.96 by MW-ICA based SVM. Efficiency and accuracy of MW-ICA based SVM was found to be very high in normal and abnormal classification.

5.5.1.3 Qualitative analysis

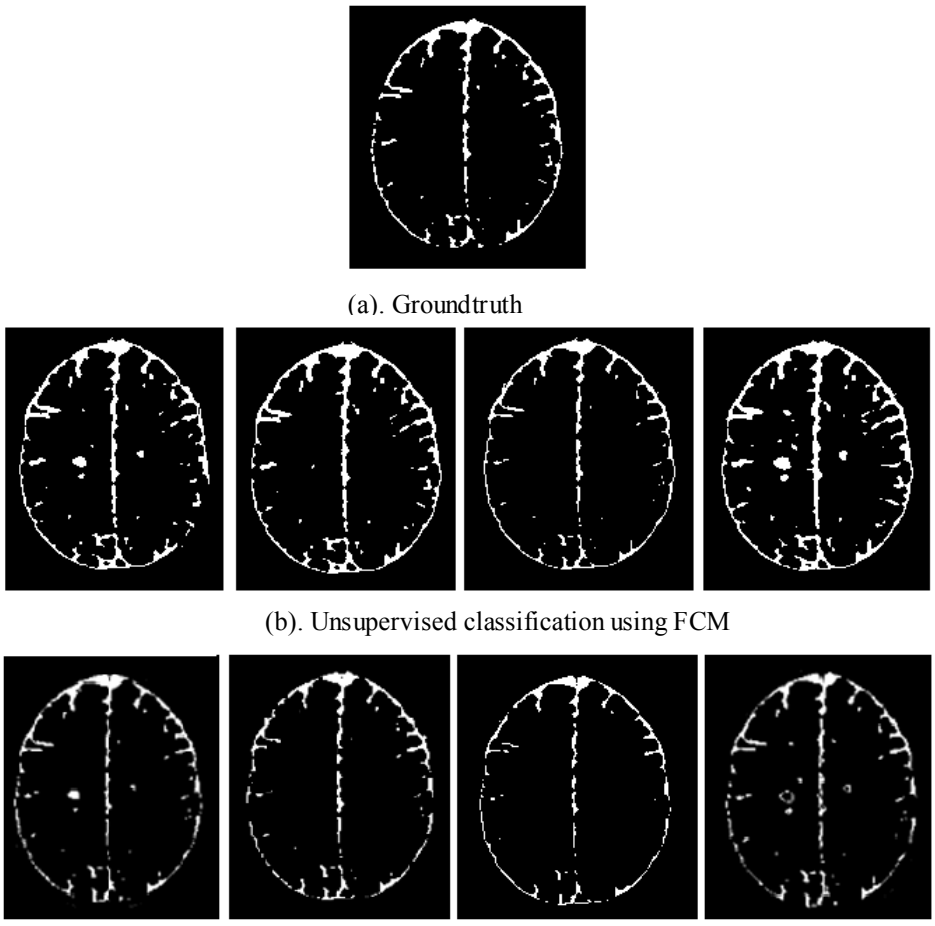
A detailed analysis of each brain tissue using groundtruth, segmented results from FCM, and classified results by SVM, based on proposed ICA

extensions and conventional ICA are provided in the following discussions. Classified results are presented after removal of extrameningial tissues, to do a better comparison with the groundtruth.

First, we consider the analysis of CSF, which is given in Fig. 5.10. Groundtruth image is shown in Fig.5.10 (a) as a gold standard for result comparison. Classified results in Fig. 5.10(b) and Fig. 5.10(c) are arranged in the order of results by SC-ICA_.06, MICA, MW-ICA and ICA from left to right. Results from MW-ICA are observed as having best similarity with the groundtruth for FCM based segmentation (Fig.5.10 (b)).

Traces of negative pixels are observed in results from all other methods. MICA+SVM also found to be very similar to groundtruth. Supporting the results in quantitative analysis, high quality results were observed from proposed SVM classifications (Fig 5.10(c)). SC-ICA_.06+SVM showed some unwanted pixels (WML) also with the true positives. Classified results from ICA, shown in the last column of Fig. 5.10 (b) and (c) proved the inefficiency of ICA in identifying local and global features from dominating background. However, ICA+SVM was found to be yielding more accurate results than ICA+FCM.

Classified GM results are available for visual analysis in Fig. 5.11 along with the groundtruth image. Results are arranged in the order of classification based on SC-ICA_06, MICA, MW-ICA and ICA from left to right. All except MICA+SVM provided classified results with more false positives. From the input image itself, the GM samples are not described as crisp values.



(a). Groundtruth

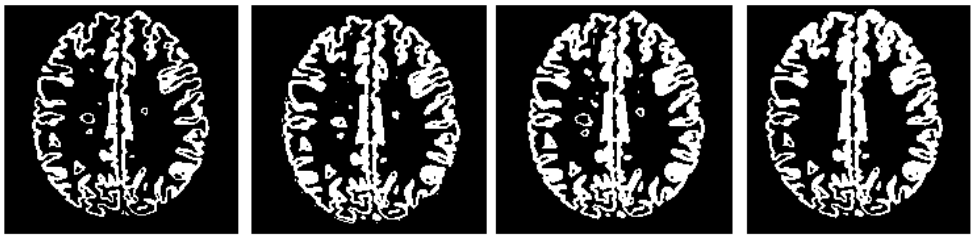
(b). Unsupervised classification using FCM

(c). Supervised classification using SVM

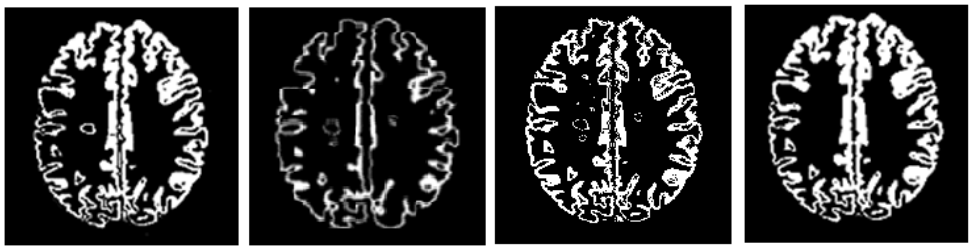
Fig. 5.10 Qualitative analysis of CSF using ground truth and classified results. Classified results from SC-ICA_.06, MICA, MW-ICA and ICA are arranged from left to right.

ICA was found to be giving poor quality results. However, traces of WML were observed in results from other methods. SC-ICA_.06 based classification could result in a better representation of GM with less false positive pixels. MICA+FCM gave comparable results, whereas MICA+SVM failed to classify some true positives. MW-ICA+SVM generated the result somewhat similar to groundtruth, but it classified some unwanted pixels also.

(a) Groundtruth



(b) Unsupervised classification using FCM

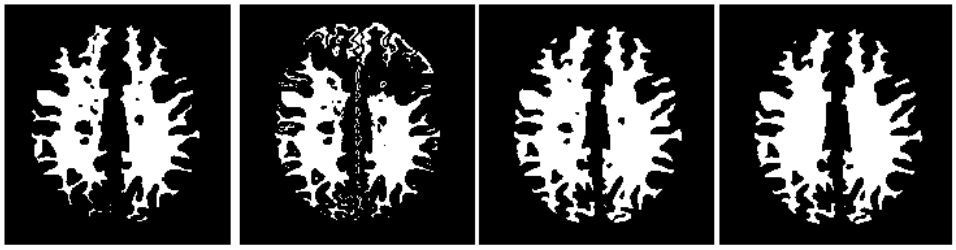


(c) Supervised classification using SVM

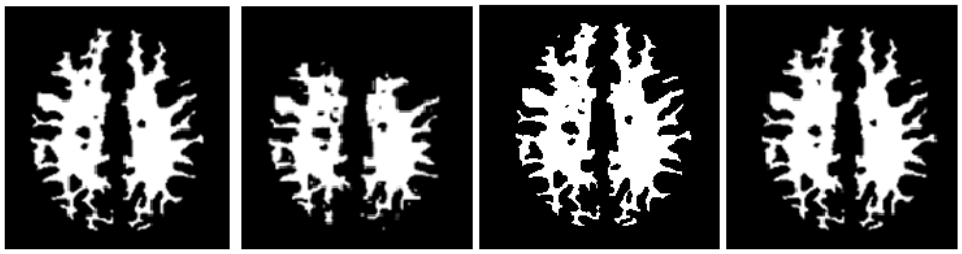
Fig. 5.11 Qualitative analysis of GM using ground truth and classified results. Classified results from SC-ICA_.06, MICA, MW-ICA and ICA are arranged from left to right.

In the Brainweb abnormal data, WM is the brain tissue affected by hyper intense lesions. Accurate identification of these small lesions and their effect on normal brain tissue analysis has great significance in this thesis work, as a part of evaluating ICA and other methods for local feature extraction. The classified WM results from ICA based methods are shown in Fig. 5.12.

(a) Groundtruth



(b) Unsupervised classification using FCM



(c) Supervised classification using SVM

Fig. 5.12 Qualitative analysis of WM using ground truth and classified results. Classified results from SC-ICA_.06, MICA, MW-ICA and ICA are arranged from left to right.

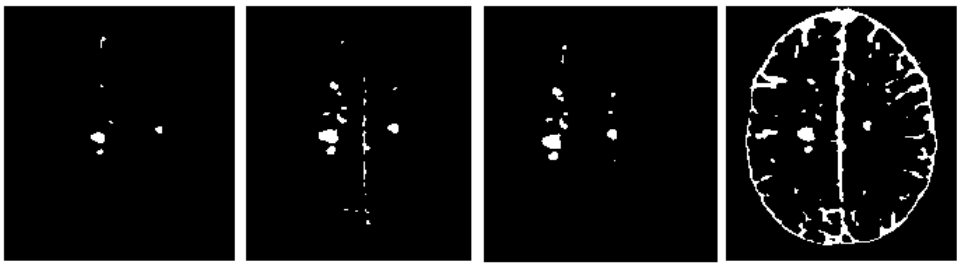
Groundtruth in Fig. 5.12(a) can be used as a reference to compare the results. When we consider the similarity of the results with the groundtruth, it is evident that MW-ICA+SVM gave best quality WM. SC-ICA based classification was also found to be good. SC-ICA+FCM identified all lesion locations, but some WM pixels also were misclassified as WML. SC-ICA+SVM identified less number of lesions with respect to goundtruth, but comparatively better than those provided by ICA based classifications. It is

noted from Fig.5.12 that ICA+FCM failed to identify the presence of abnormalities in WM, whereas classifications based on MICA failed to correctly identify all the WM pixels in their results. MICA+FCM results showed highest misclassifications (Fig. 5.12(c) 2nd column). The effect of global feature suppression in MICA algorithm is very clear from these results. However, the abnormality positions were correctly identified. MW-ICA+FCM could not give the performance of MW-ICA+SVM, even though it performs better than those from ICA based classifications.

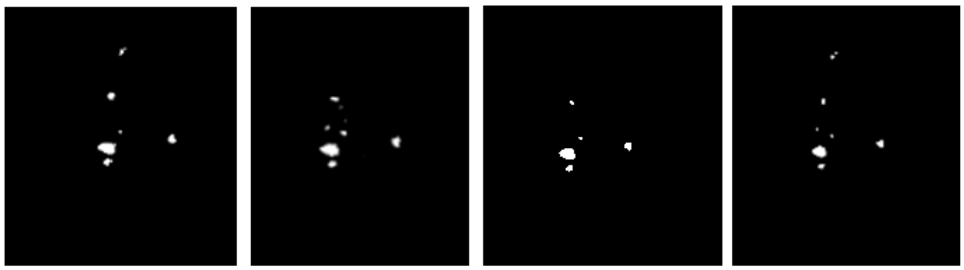
One of the main objectives of this thesis is to resolve the issue in ICA for small abnormality analysis. The effectiveness of the proposed methods in achieving this can be best described by small WML analysis. Important observations from Fig. 5.13 are,

- In ICA based FCM results, CSF and WML were found to be difficult to separate. However, ICA+SVM provided a better performance, because of the supervised learning process with best WML features.
- SC-ICA based classification results were found to be showing good accuracy in WML detection, but available information was comparatively less than that from groundtruth.
- MICA+FCM identified all the WML locations, paying the cost of a large number of misclassifications also. But MICA+SVM was found to be yielding accurate results, even though it failed to identify all the WML locations.
- MW-ICA+FCM also performed very well in WML classification, but additional misclassified pixels affected the quality of the segmented images. MW-ICA+SVM solved this issue with more accurate and reliable classified lesions.

(a) Groundtruth



(b) Unsupervised classification using FCM



(c) Supervised classification using SVM

Fig. 5.13 Qualitative analysis of WML (abnormality) using ground truth and classified results. Classified results from SC-ICA_.06, MICA, MW-ICA and ICA are arranged from left to right.

Qualitative and quantitative analysis results in this section recommended SVM based on proposed extensions as promising methods in brain tissue classification. MW-ICA+FCM was observed to be yielding high performance in almost all results. A more detailed study in clinical analysis can reveal the true behavior of these approaches with clinical MRI analysis.

5.5.2 Clinical image analysis

Unsupervised classification by FCM (experiment 2), supervised analysis by SVM (experiment 3) and qualitative analysis (experiment 4) were repeated for clinical dataset also in the same environment as discussed in the previous section. Both normal and abnormal datasets were considered in the performance evaluation. Experiments followed the validation and test plan given in Table 5.1. Abnormal datasets were divided into two categories; hyper-intensity lesions and other abnormalities. As a first step, experiment 1 was performed with SC-ICA for various α values, MICA and MW-ICA for different daubechies wavelets, and conventional ICA. We considered five classes, CSF, GM, WM, hyper intense lesions (WML) and other abnormalities (tumors, infarcts, edema etc.), for brain tissue classification. Performance improvement by proposed methods over ICA based feature extraction in unsupervised and supervised classification was studied with the help of FCM and SVM respectively.

5.5.2.1 Unsupervised classification results

In the unsupervised analysis of clinical images, we considered only the normal and abnormal test data described in Table 5.1, to generate comparable results with SVM based classifications. FCM with default parameter setting was applied to generated results from SC-ICA for thresholds 0.12, 0.1, 0.6 and 0.3. The segmented results were labeled and classified with the help of an experienced radiologist. The same steps were repeated for results from MICA and MW-ICA for *db6* wavelet, and conventional ICA also. The observed results are presented in Table 5.5, and in Fig. 5.13 (a) and Fig. 5.13 (b). A visual analysis is also presented with the help of a clinical case with WML. Average Tanimoto Index values from 48 normal cases and 64 abnormal cases are tabulated in Table 5.5. For SC-ICA, normal case analysis was done with α value, 0.1, 0.12 and 0.03. In abnormal case we reduced the

threshold a little such as 0.1, 0.06, and 0.03 to work with small abnormalities. According to the spectral angle calculation discussed in Chapter 4, we selected a value > 0.06. Threshold value 0.03 is also included to explain the effect of more clusters in normal and abnormal analysis.

Table 5.5 Tanimoto Index values from clinical analysis

		Normal				Abnormal				
Methods	α	CSF	GM	WM		α	CSF	GM	WM	Abn.
SC-ICA+FCM	.1	0.74	0.64	0.81		.1	0.68	0.55	0.78	0.78
	.12	**0.90**	**0.69**	**0.89**		**.06**	**0.81**	**0.78**	**0.88**	**0.89**
	.03	0.72	0.60	0.63		.03	0.79	0.33	0.89	0.90
MICA+ FCM		0.68	0.64	0.82			0.70	0.78	0.81	0.92
MW-ICA+FCM		**0.84**	**0.78**	**0.88**			**0.79**	**0.82**	**0.84**	**0.91**
ICA + FCM		0.82	0.42	0.78			0.75	0.31	0.80	0.76

Abn. = Abnormality

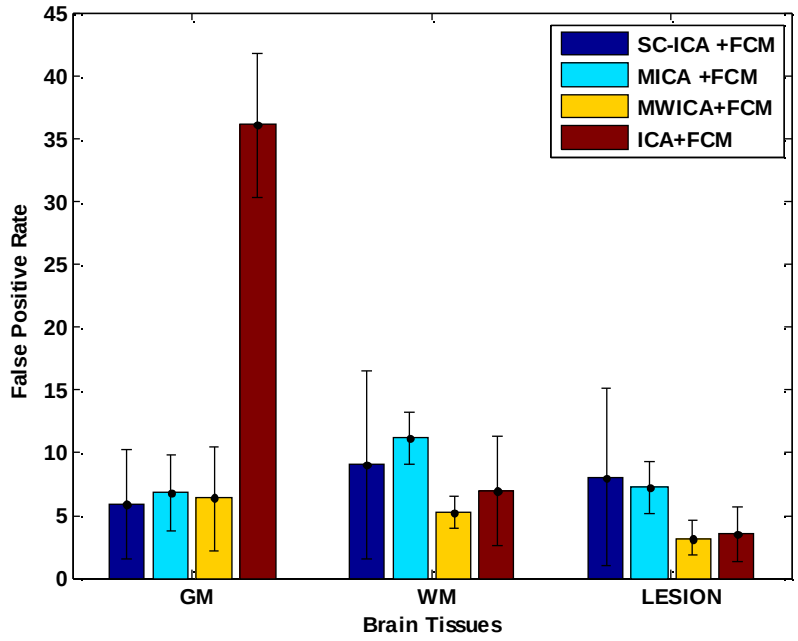
FCM based on proposed methods were found to be giving very good results for almost all tissue classes in this study. In normal case analysis, SC-ICA based FCM with threshold value 0.12 was observed as yielding best TI values for CSF/WM, 0.90/0.89, whereas MW-ICA+FCM also showed similar performance with best TI for GM, 0.78. SC-ICA based classification for lower thresholds and MICA+FCM failed to provide a better performance than ICA based analysis for some tissues like CSF. In abnormal data analysis, more than three tissue classes were to be extracted from three input images. Almost similar behavior as observed in normal case is provided in the classification results by proposed and conventional methods.

SC-ICA_.06+FCM and MW-ICA+FCM were observed as yielding very good TI values. For WM and abnormal tissue classification, SC-ICA for

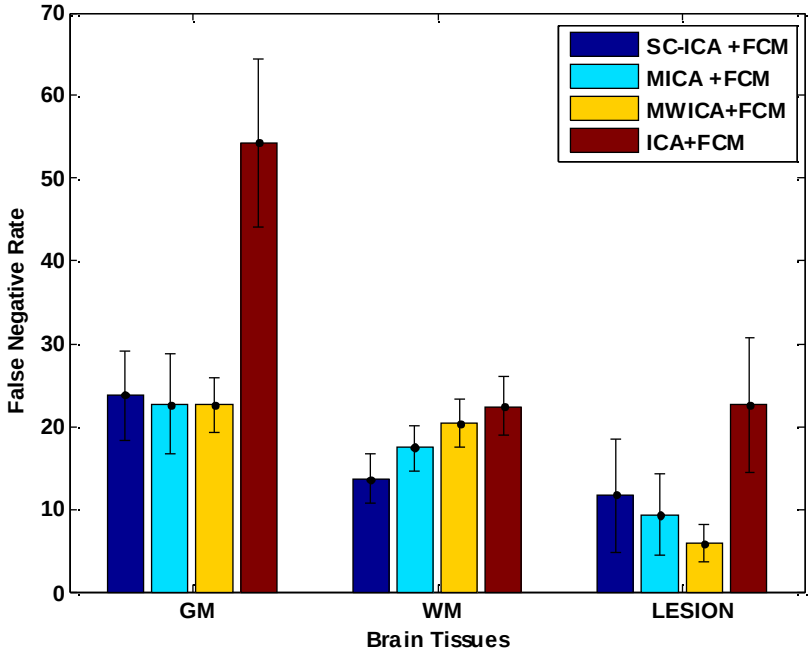
threshold value 0.03 showed high TI values. More clusters generated as a result of reduction in threshold improved the classification of WML and other abnormalities. MICA+FCM provided the best average TI, 0.92, for abnormality detection. However, as it is mentioned in the MICA algorithm, global tissues were found to be suppressed in the results.

Cost of predicting a negative tissue as a positive, or positive tissue as negative is very high in MRI analysis. False Positive Rate (FPR) and False Negative Rate (FNR) observed for 35 abnormal cases showing hyper intense lesions by four methods under consideration were used in this study to analyze these misclassifications. Wide variation in measures observed for different cases. So average of estimated results with standard deviations is shown as error bars in Fig. 5.14(a) and (b) for GM, WM and Lesions. Fig. 5.14(a) shows that GM classification from MICA and MW-ICA has almost same FPR. SC-ICA+FCM shows the least average FPR, but with wide variations.

In Fig. 5.14(b), ICA+FCM was observed as yielding highest FNR among all the classification approaches, whereas MW-ICA+FCM was observed as best classifier for CSF and abnormalities with least values for FNR and variations. However SC-ICA+FCM turned as the best method yielding least FNR with least standard deviations for WM classification. The significant reduction in FPR and FNR, observed for MW-ICA based classification, is a promising result in clinical analysis. ICA+FCM showed the maximum FPR for GM. However, in the case of WM and abnormality analysis, comparatively lower values observed for MW-ICA+FCM and ICA+FCM based analysis. MW-ICA+FCM was found to be giving least averages with least standard deviations.



(a). FPR comparison of segmented tissues



(b). FNR comparison of segmented tissues

Fig. 5.14 Misclassification rate analysis of hyper intense lesions

A subjective analysis of these observations was performed in the qualitative study with an abnormal case. Fig. 5.15 shows input slices selected for the analysis, in the order of T1WI, T2WI and FLAIR image from left to right. The selected case shows periventricular ischaemic change with few chronic lacunar infarcts in both cerebral hemispheres as hyper intense lesions. WM details are available from T1WI; T2WI shows CSF and abnormal points. FLAIR images give information on lacunar infarcts (hyper intense). Not much detail on GM was available from these images.

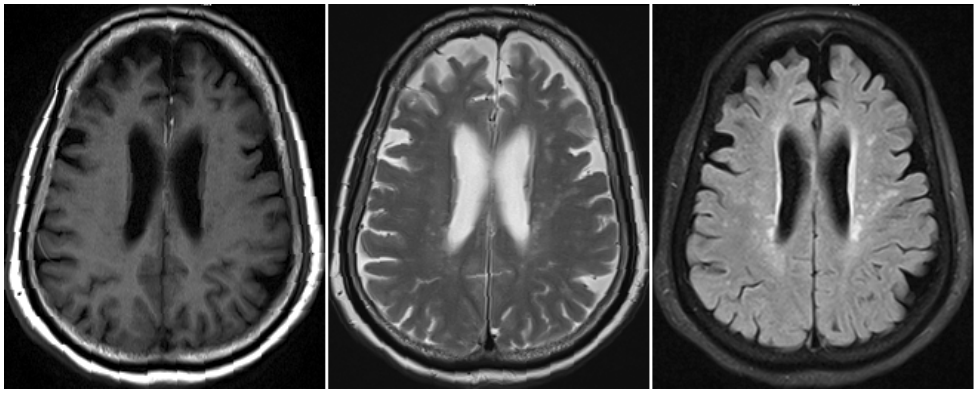
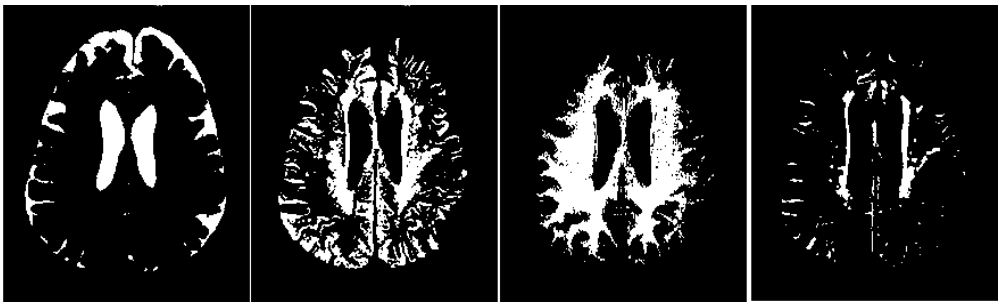
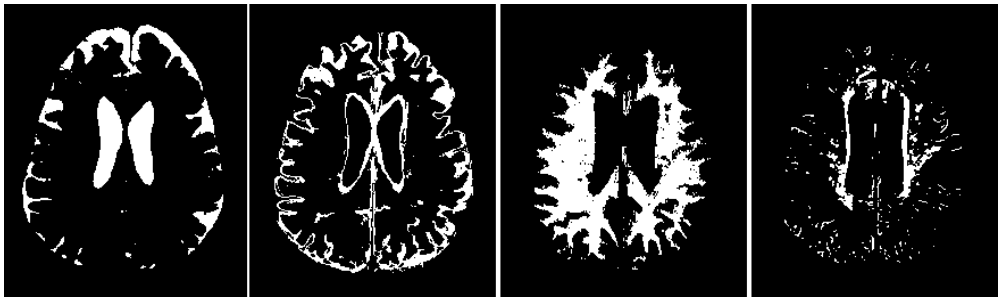


Fig. 5.15 Clinical input images, T1WI, T2WI and FLAIR from left to right.

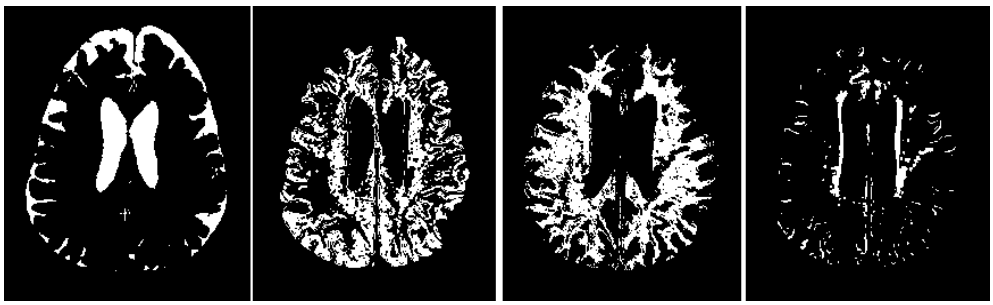
ICA based FCM was first applied on this dataset. Classified results are shown in Fig. 5.16 (a). Then SC-ICA for threshold 0.06 was executed with this input dataset, from which FCM generated the results given in Fig. 5.16 (b). We repeated FCM on results from MICA and MW-ICA for *db6* wavelet using the same dataset. Segmented brain tissues, CSF, GM, WM and lesions are shown in Fig. 5.16 (c) and Fig. 5.16 (d) respectively. Column wise comparison of Fig. 5.16 gives a detailed picture of the improvement in tissue classification by the proposed methods, which can be highlighted as follows.



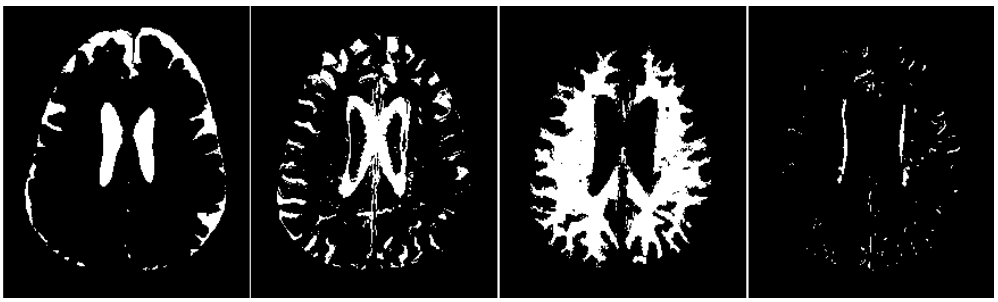
(a) ICA based FCM; CSF, GM,WM and WML from left to right



(b) SC-ICA+FCM; CSF, GM, WM and WML from left to right



(c) MICA+FCM; CSF , GM,WM and WML from left to right



(d) MW-ICA+FCM; CSF, GM,WM and WML from left to right

Fig. 5.16 Unsupervised brain tissue classification from a clinical case

- CSF details are found to be better represented by SC-ICA+FCM and MICA+FCM in Fig. 5.16 1st column.
- GM (Fig. 5.16 2nd column) is found to be better in SC-ICA+FCM and MW-ICA+FCM; MW-ICA has shown a few unwanted pixels due to misclassifications.
- In input T1WI, lesion details are hypo intense, which cannot be distinguished from WM. ICA+FCM also failed to locate these points in WM. SC-ICA+FCM has shown the best results in this regard. Comparable results are observed from MW-ICA+FCM also. MICA+FCM represented them with a penalty of losing true positives in the result (Fig. 5.16 3rd column).
- For WML segmentation (Fig. 5.16 4th column), MW-ICA+FCM performed best with accurate representation of abnormalities and reduced presence of negative pixels. ICA results in Fig. 5.16 (a) failed to give specific details of abnormality. SC-ICA+FCM and MICA+FCM improved the results; still the presence of unwanted pixels affects the overall quality of the results.

Quantitative and visual analysis demonstrated the efficiency of the proposed methods in improving brain tissue classification. SC-ICA+FCM and MW-ICA+FCM presented a very good performance in identification of small abnormalities. As a last step, we conducted a supervised learning, classification and analysis of clinical dataset as discussed in the next section.

5.5.2.2 Supervised classification results

From synthetic image analysis, it was observed that supervised classification is providing more accurate results than FCM. To ensure the positive effect of the supervised approaches with proposed methods in brain tissue classification, experiment 3 was conducted with SVM. Supervised learning for normal and abnormal data was conducted with features extracted from

multispectral slice sets, as described in validation plan in Table 5.1. Feature vectors collected from training data were utilized to generate the classification model for brain tissue analysis. The performance measures sensitivity, specificity and accuracy were estimated separately for each healthy and non healthy case. Three multispectral slice sets were considered for each case to collect the feature vectors. Performance measures observed from different cases were found to be widely varying. Average and standard deviation of the estimated values are tabulated in Table 5.6. GM, WM and abnormality are included in the analysis. Both hyper intense lesions and other abnormalities were merged into the broad category '*Abnormality*'.

Table 5.6 Performance analysis of classified tissues from supervised approach

		GM			WM			Abnormality		
Methods	α	Spec. ±std.	Sens. ±std.	Acc. ±std.	Spec. ±std.	Sens. ±std.	Acc. ±std.	Spec. ±std.	Sens. ±std.	Acc. ±std.
SC-ICA +SVM	.1	68.32 ±5.44	80.29 ±5.65	70.86 ±6.95	91.15 ±5.69	90.32 ±6.97	92.20 ±4.39	93.99 ±5.11	46.76 ±16.97	78.30 ±12.02
	.06	94.17 ±5.36	58.91 ±10.62	76.40 ±5.43	91.01 ±8.50	84.81 ±6.18	90.78 ±6.07	91.99 ±7.43	57.19 ±21.47	82.31 ±14.9
	.03	84.25 ±5.37	56.63 ±11.53	75.43 ±5.19	90.18 ±4.24	90.34 ±5.06	91.68 ±4.08	92.77 ±7.12	90.28 ±5.59	92.57 ±6.25
MICA +SVM		93.18 ±3.01	80.71 ±4.34	89.81 ±2.17	88.91 ±2.10	80.76 ±4.08	85.09 ±3.74	95.26 ±2.11	63.35 ±3.32	89.78 ±2.27
MW-ICA+SVM		**93.68 ±4.20**	**83.32 ±6.12**	**90.61 ±3.04**	**94.83 ±1.28**	**88.95 ±7.14**	**93.60 ±2.18**	**96.88 ±1.40**	**91.99 ±3.88**	**95.69 ±1.38**
ICA + SVM		63.92 ±5.73	57.62 ±4.67	62.52 ±4.55	94.08 ±4.36	85.50 ±8.20	92.37 ±2.26	97.57 ±2.22	84.90 ±5.02	89.31 ±2.18

Spec. = Specificity, Sens. = Sensitivity, Acc. = Accuracy, std. = standard deviation

Main observations from Table 5.6 can be summarized as follows,

- Considering GM classification results, best average results with least standard deviations were provided by MW-ICA+SVM and MICA+SVM. MW-ICA+SVM gave an accuracy of 90.61% with standard deviation 3.04% and MICA+SVM showed 89.81% accuracy with 2.17% variation.
- In WM classification also, MW-ICA+SVM presented the best performance with specificity/sensitivity/accuracy, 94.83/88.95/93.6, and consistency with decreased standard deviations. SC-ICA+SVM results were also found to be very good, but poor performance was observed for MICA+SVM.
- In the case of abnormality analysis, we considered small as well as large type of lesions and tumors. MW-ICA+SVM was observed as the best classifier with consistent results 96.88/91.99/95.69 as specificity/sensitivity/accuracy values. SC-ICA+SVM for threshold 0.03 also provided a high performance, but estimated standard deviation values indicated the inconsistency in results due to large inter-case result variations.

The classification models obtained from the validation process were utilized to segment the brain tissues using the test cases given in Table 5.1. The classified tissues were compared with manually segmented images to measure the Tanimoto Index (TI), and average of the observed values for normal and abnormal cases excluding WML cases are summarized in Table 5.7. SC-ICA was performed with threshold values 0.12, 0.1, and 0.03 for normal case, and abnormal case analysis was conducted with threshold values, 0.1, 0.06 and 0.03. SC-ICA_0.12+SVM was observed as the best performer for normal case analysis with highest TI value for CSF/WM, 0.88/0.87, and high TI value, 0.66, for GM. ICA_0.06+SVM founds to be yielding best values for abnormal case analysis for CSF/GM/WM,

0.77/0.78/0.90. However, abnormality classification by MW-ICA+SVM provided highest TI value, 0.91. ICA+SVM failed to provide competitive results for all cases except WM. To demonstrate the effect of the proposed methods in supervised classification qualitatively, a clinical case from abnormal data is analyzed in depth. Abnormality in clinical case shown in Fig. 5.17 was observed as a lesion surrounded by edema. T2WI shows the location of abnormal tissues as CSF-like hyper intense, but not distinguishing the tumor and edema.

Table 5.7 Tanimoto Index from classified tissues in clinical analysis

		Normal				Abnormal				
Methods	α	CSF	GM	WM		α	CSF	GM	WM	Abn.
SC-ICA + SVM	.1	0.79	0.63	0.83		.1	0.43	0.53	0.88	0.78
	.12	**0.88**	**0.66**	**0.87**		**.06**	**0.77**	**0.78**	**0.90**	**0.79**
	.03	0.60	0.57	0.51		.03	0.63	0.29	0.89	0.75
MW-ICA + SVM		0.63	0.68	0.84			0.62	0.68	0.82	0.91
MICA+SVM		0.64	0.66	0.80			0.60	0.69	0.74	0.75
ICA + SVM		0.61	0.45	0.77			0.41	0.47	0.84	0.62

Abn. = Abnormality

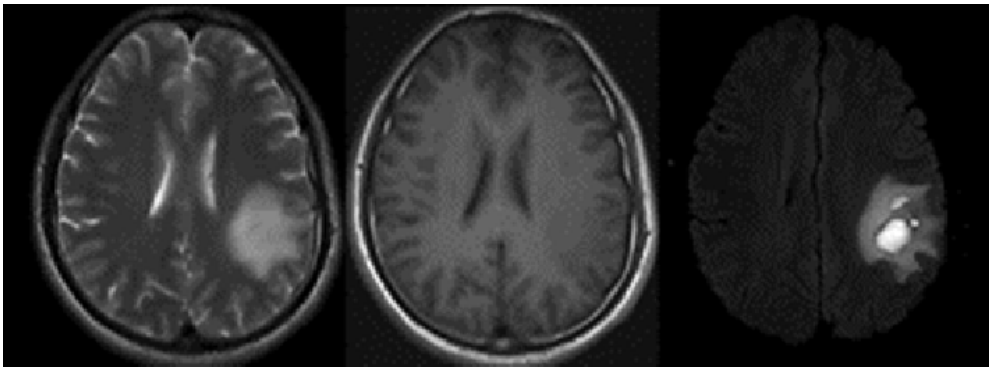
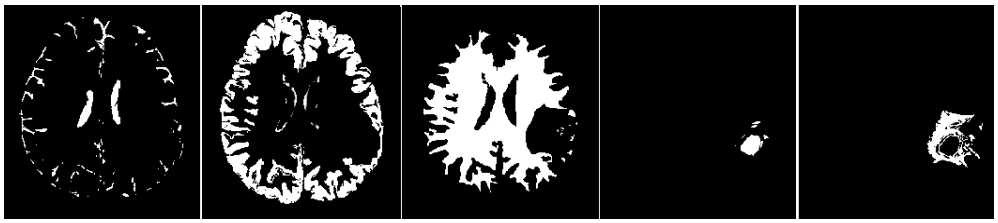


Fig. 5.17 Clinical input images for SVM analysis, T2WI, T1WI and FLAIR from left to right.

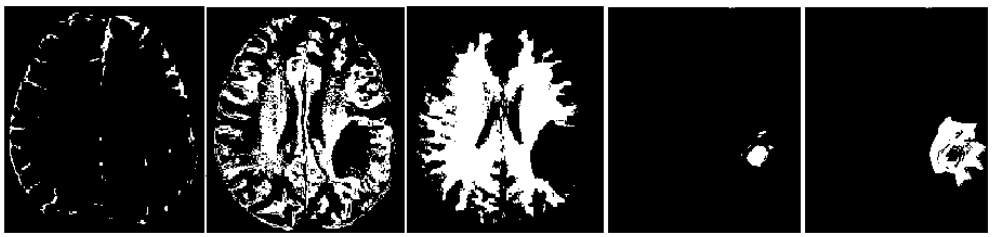


(a) Manual segmentation

(b) ICA+SVM; CSF, GM, WM and tumor + edema from left to right

(c) SC-ICA_06+SVM; CSF,GM,WM and tumor + edema from left to right

(d) MICA+SVM; CSF, GM, WM and tumor + edema from left to right



(e) MW-ICA+SVM; CSF, GM, WM tumor, surrounding edema from left to right

Fig. 5.18 Results from manual segmentation and supervised classification

T1WI also showed the similar behavior, locating abnormality with vague hypo intense part. FLAIR image well described the tumor and edema, but failed to incorporate the information of other tissues in it. Fig. 5.18 describes the results from proposed SVM classifications and conventional ICA based classification. Manually segmented tissues shown in Fig. 5.18(a) were used as the reference for comparison. From the qualitative analysis of results in Fig 5.18, following points can be highlighted.

- ICA+SVM results in Fig. 5.18(b) best described the lesion part, but failed to identify the surrounding edema. However, high quality results were observed for WM, exactly removing the affected portions. It could not show a high quality performance in classification of CSF and GM.
- Classification results from SC-ICA_.06+SVM (Fig. 5.18 (c)) were found to be more promising with high quality results for all tissues except GM. SC-ICA+SVM results specifically located the abnormality, and its effect on other tissues in a better way, which seems to be very useful in clinical analysis.
- MICA+SVM results showed the lesion and the surrounding edema (Fig. 5.18 (d) last column) with a clear description of the separation between lesion and edema. However, MICA cannot reach the performance of SC-ICA or ICA in classification of GM, WM and CSF.
- As a method providing solution to over-complete issue in ICA, MW-ICA could separate the edema and tumor into two results as given in Fig. 5.18 (e) (last two columns). High quality CSF and WM were observed, with a comparable accuracy to the manually segmented results. A higher rate of misclassifications observed for GM, leading to a poor quality result as in Fig. 5.18(e) 2nd column.

As a concluding remark it can be stated that supervised classification results were found to be superior to unsupervised analysis results. Classifications

based on MW-ICA and SC-ICA provided high performance classification for almost all brain tissues. However, SC-ICA was found to be yielding varying results and performance for different threshold values. Threshold selection associated with SC-ICA, which is highly dependent on the data characteristics, was observed as a great challenge in MRI analysis. MICA based classification results were found to be very good for abnormality analysis. However, global feature suppression in MICA algorithm decreased the classification performance of other tissues, CSF, GM, and WM. MW-ICA and MICA based classification results were found to be slightly varying on different wavelets. In general, low order daubechies wavelets gave good results.

5.6 Summary

Quantitative and qualitative analysis of the proposed ICA extensions and conventional ICA, with supervised and unsupervised classification approaches is discussed in this chapter. A detailed description on the datasets used, preliminary steps to prepare the data for feature extraction are described in the initial sections. Image registration and feature collection are described with the help of examples. Bias estimation for different validation techniques also included as a part of the feasibility study. Result analysis of synthetic and clinical datasets is divided into two, quantitative and qualitative analysis. Supervised classification using SVM, and unsupervised classification using FCM were performed on normal and abnormal multispectral sets in synthetic and clinical database. Compared to conventional ICA based classifications, proposed methods have shown significant improvement in brain tissue classification. Potential of these methods in improving the results, especially in the analysis of small abnormalities are found to be very promising in pathological study and disease progress evaluation.

Chapter - 6

A HYBRID APPROACH TO BRAIN TISSUE CLASSIFICATION

6.1 Introduction

Chapter 4 discussed the proposed ICA extensions, and Chapter 5 evaluated the supervised and unsupervised classifications based on these methods using synthetic and clinical data. In this chapter, we propose a new hybrid feature selection method for supervised brain tissue classification using SVM. ICA and multisignal wavelet analysis in the spectral domain are the techniques used in the feature extraction phase. For each tissue sample, corresponding values in independent components and wavelet coefficients were simultaneously selected for training and classification. Section 6.2 explains the proposed method. Both synthetic and clinical datasets were included in the performance analysis of the proposed algorithm. Details of the experiment are given in Section 6.3. The observed results are summarized in Section 6.4. Selected classification models with optimal performance were used in the qualitative analysis of each tissue in section 6.5.

6.2 Proposed method for SVM classification

6.2.1 Method

Wavelet decomposition of the input spectra and ICA are the core concepts used in this algorithm. Major steps involved in this method are depicted in Fig. 6.1. Co-registered images from different MRI sequences were collected to form a multispectral suite. Each pixel vector represents the spectral signature of the area specified by that pixel, and collection of these spectral signatures generated multisignals, as discussed in Chapter 4. Two

independent operations (can be implemented as parallel) are applied on these multisignals to generate the feature sets; Conventional ICA using FASTICA as discussed in Chapter 3, and 1-D multisignal wavelet analysis, as shown in Fig 6.1. Optimal features for supervised learning are collected from the generated source components (basis vectors) from these methods, with the help of an experienced radiologist. Non-linear SVM using RBF kernel with default settings, as described in Chapter 5, is applied on the training data and test data for brain tissue classification. Details of the feature selection are discussed with the help of an example in the following sections.

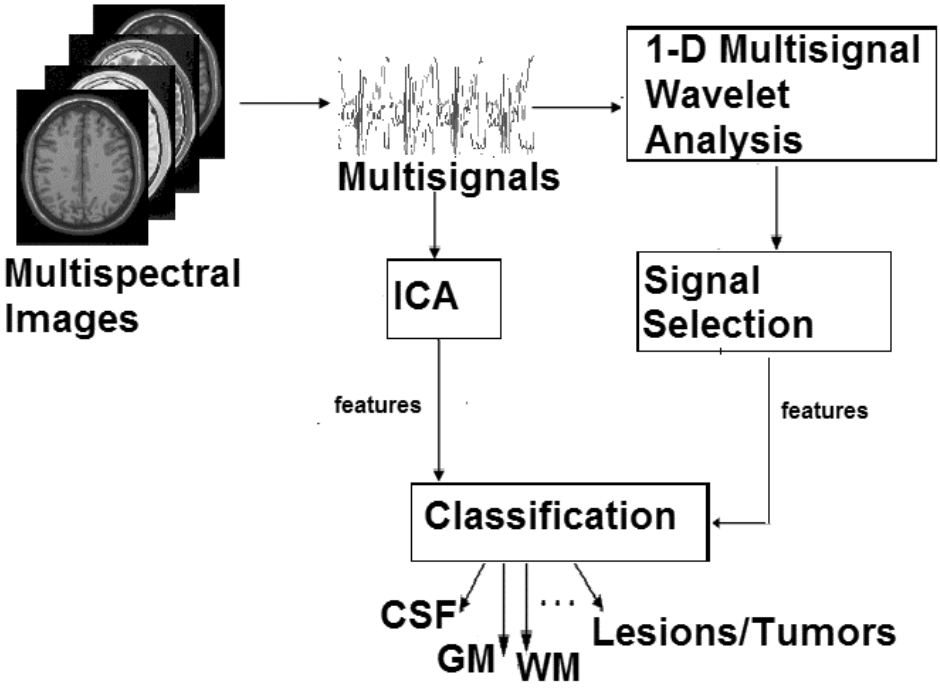


Fig. 6.1 Proposed Hybrid SVM classification

6.2.2 Signal selection from wavelet analysis

Discrete wavelet transform is applied on input image cube in the spectral domain as discussed in Chapter 4, exploiting the same concept used in the spatial domain [53, 55]. In multisignal wavelet analysis, approximation and detail coefficients at different levels of decomposition yields different

decorrelated basis components. Therefore, the problem with wavelet decomposition is to decide the level of decomposition [55] for selection of best features. A simple method using the correlation coefficient, ρ, between detail coefficients at each level helps to do this. The magnitude of the computed positive correlation coefficient shows the degree of similarity between the subcomponent images. Correlation coefficient 'ρ' is given by,

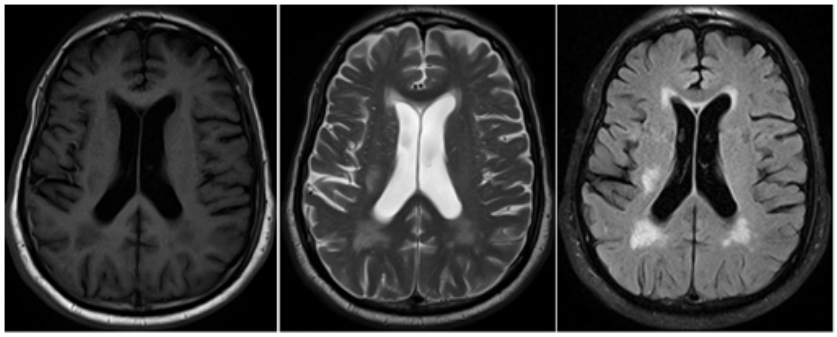
$$\rho(D_1, D_2) = \frac{\text{cov}(D_1, D_2)}{\sigma_{D_1} \cdot \sigma_{D_2}} \tag{6.1}$$

where 'cov(D_1, D_2)' is the covariance between datasets 'D_1' and 'D_2', σ_{D_1} and σ_{D_2} are the standard deviations of 'D_1' and 'D_2' respectively.

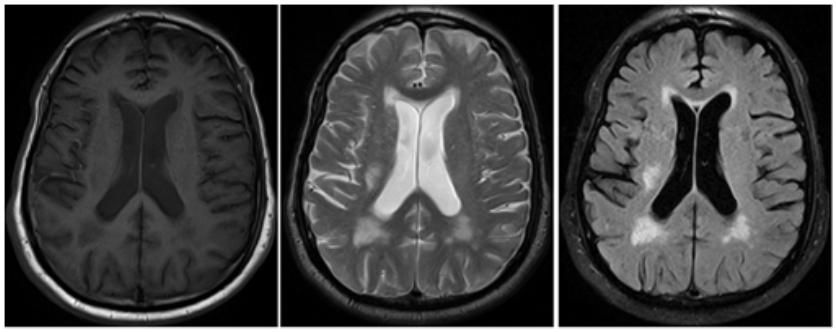
For example, consider the correlation analysis to select the best level of wavelet decomposition for feature selection using a clinical case consisting of T1WI, T2WI and FLAIR images, as given in Fig. 6.2(a). Three levels of decomposition are considered. Correlation matrix, CM, is observed as,

CM =			CM =		
Level 1	ρ(1,2)	ρ(2,3) ρ(1,3)	0.8283	0.4919	0.0730
Level2	ρ(1,2)	ρ(2,3) ρ(1,3) ⟹	-0.9047	-0.7865	-0.8269
Level3	ρ(1,2)	ρ(2,3) ρ(1,3)	-0.9881	-0.9246	-0.6213

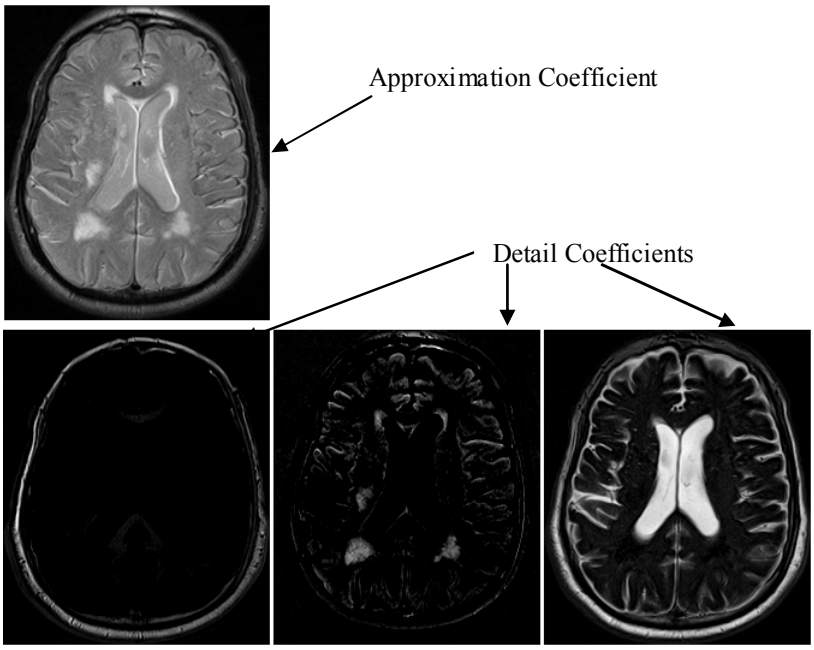
Minimum correlation between subband components are provided by the column wise minimum of CM. CM_{min}= [-0.9881 -0.9246 -0.8269] with Index Vector = [3 3 2]. So components 1&2, and 2&3 have minimum correlation at level 3. Images with negative correlation are also considered as dissimilar. So index value 3 is considered as the best level of decomposition in this example to select the candidates for feature selection from detail coefficients and approximation coefficient. Independent components are given in Fig 6.2 (b) and selected wavelet coefficients are shown in Fig. 6.2 (c).



(a) Input images



(b) Independent Components



(c) Wavelet coefficients for level 3

Fig. 6.2 Feature extraction using ICA and multisignal wavelet analysis

6.2.3 Feature selection

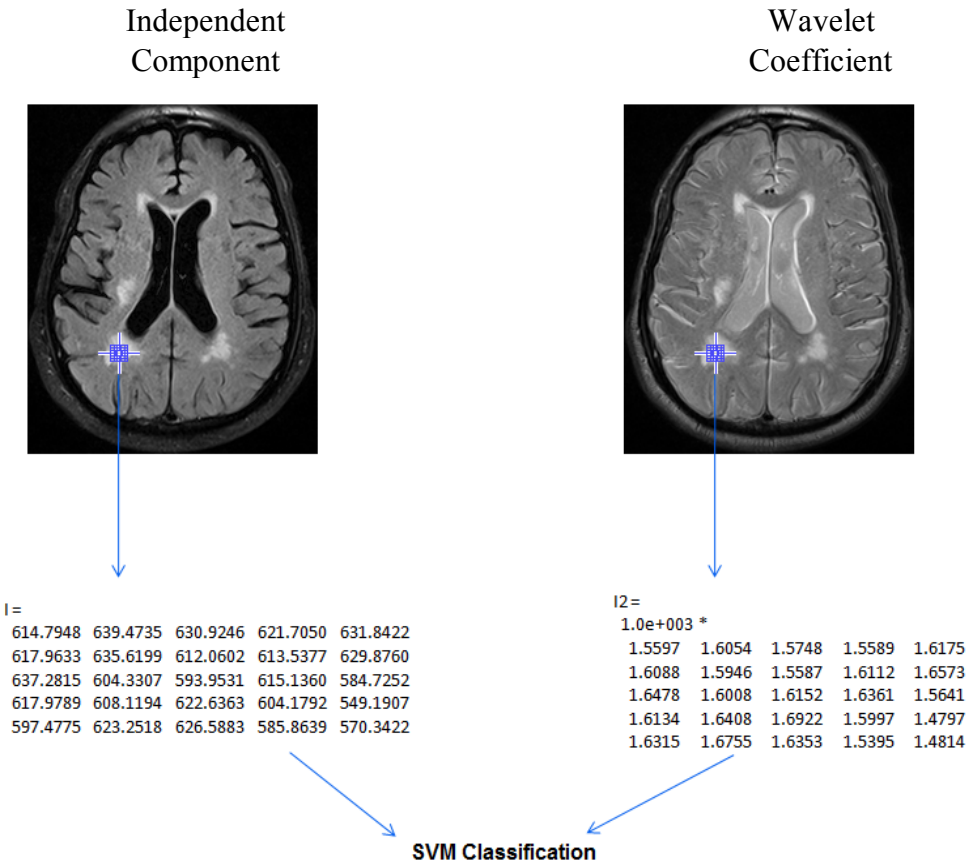


Fig 6.3 Feature selection from Independent Component and wavelet coefficient

Both ICs and wavelet coefficients give features of source components or basis vectors. Feature selection for white matter lesion is shown in Fig 6.3. A 5x5 window is chosen to locate the optimal feature set. For each sample, two features are considered; feature from independent component, and corresponding value from wavelet coefficient. Since it is a pixel based classification, it can be used in segmentation also. The proposed hybrid approach to SVM classification is summarized in Algorithm 6.1. Step2a and Step2b are independent, and can be implemented and performed in parallel to reduce the computational complexity.

Input: Co-registered slices from MR sequences, T1WI, T2WI, PDI, FLAIR etc.

Step1: Consider input multispectral image as a collection of spectral signatures. Apply 1-D multisignal wavelet decomposition on these signals to divide the spectral domain into low frequency and high frequency components.

Step 2: Apply two parallel operations for feature extraction as follows,

Step 2a: Calculate the correlation coefficient between the high frequency subcomponents for each level using the eq.,

$$\rho(D_1, D_2) = \frac{\mathrm{cov}(D_1, D_2)}{\sigma_{D_1} \cdot \sigma_{D_2}}$$

where $\mathrm{cov}(D_1, D_2)$ is the covariance between datasets 'D_1' and 'D_2', and σ_{D1} and σ_{D2} are the standard deviations of 'D_1' and 'D_2' respectively. Select the wavelet coefficients providing minimum degree of correlation for feature selection.

Step 2b: Apply FASTICA algorithm on input dataset to generate independent components.

Step 3: Select Independent component features and corresponding wavelet features using a 3x3 or 5x5 window.

Step 4: Apply SVM training and classification on selected features to classify the brain tissues.

Output: Classified/segmented brain tissues and abnormalities.

Algorithm 6.1 The proposed hybrid feature selection method for SVM classification

6.3 MR images and experimental setup

Both synthetic and clinical images were considered to evaluate the performance improvement of the proposed method. The synthetic MR image

analysis includes a dataset containing Multiple Sclerosis (MS) data obtained from the BrainWeb.

Table 6.1 Evaluation plan for classification

Category	Training data		Cross validation	Test data		Feature vectors
	Cases	Slices		Cases	Slices	
Synthetic normal	1	40	ECV 10	1	20	1800
Synthetic abnormal	1	40		1	20	1800
Clinical normal	120	360		48	144	7560
Clinical abnormal (WML)	75	225		35	105	6600
Clinical abnormal (other)	75	225		29	87	6240

60 multispectral datasets containing axial T1WI, T2WI, and Proton Density Images (PDI) from each database (normal and abnormal) were considered for analysis. Slices from each sequence were selected with parameter settings 1-mm slice thickness and noise level of 0%. T1WI, showing WM and GM components, T2WI showing CSF details and White Matter Lesions (WML), and PD images showing more abnormal tissue details were considered in the analysis. Clinical image analysis was conducted using normal and abnormal datasets consisting of T1WI, T2WI and FLAIR sequences. MATLAB based registration, as explained in Chapter 5, was performed on these dataset to generate co-registered images.

Two sets of experiments were done using these databases. First experiment performs training and classification using SVM, and its analysis. Second experiment uses the models generated from the first set to evaluate the segmented images. Details of the experiment 1 are shown in Table 6.1. Total 382 patients were selected for clinical analysis, out of which 168 cases were

normal and 214 were abnormal. Abnormal cases were again categorized into White Matter Lesions (WML) and other abnormalities. For classification, we considered 110 WML cases and 104 other abnormal cases. Three slices, best representing the abnormal locations were collected from each sequence. CSF, GM, WM and Abnormalities were the tissues considered, and total features included in the experiment were given in Table 6.1. Validation and model selection were conducted by external 10 fold cross validation. In the second set of experiments, the classification models were considered in the segmentation of synthetic and real brain tissues. Performance evaluation was done with statistical measures sensitivity, specificity, accuracy, error bar diagrams and ROC curves [77].

6.4 Classification results

Feature extraction process using multisignal wavelet analysis and ICA were first performed to collect the best features of each tissue. Non-linear SVM (RBF kernel, $\sigma = 1$) functions in Matlab 7\Pattern Recognition Toolbox, with default parameter settings were applied on these selected features for training and validation. One-against-all SVM strategy was adopted to solve the classification problem. Sequential Minimal Optimization method was chosen to find the optimal separating hyper-plane. The same procedures were repeated for SVM coupled with ICA alone also, to give a detailed comparative analysis as discussed in the following sections.

6.4.1 Synthetic image analysis

Total 120 synthetic multispectral slice sets were considered for SVM classification for normal and abnormal databases. Distribution of the data into training and test data categories are shown in Table 6.1. ECV10 is applied on the 40 normal multispectral sets to select the best classification model. Then, a separate set of 20 slices in test data were utilized to measure

the performance measures like sensitivity, specificity and accuracy. Average of these measures is tabulated in Table 6.2. The same steps were repeated for 60 (both training and test data) abnormal multispectral sets also, and results are included in Table 6.2. To evaluate the improvement in classification performance, SVM learning and classification were performed on normal and abnormal features from ICA also, in the same environment. The observed results for CSF, WM, GM and WML supported the proposed method as an efficient method for brain tissue classification. Variations in results among different multispectral sets were negligible, since the synthetic data were of similar characteristics.

Table 6.2 Analysis of synthetic brain tissues

		Normal				Abnormal		
Tissues	Method	Sens.	Spec	Acc.		Sens.	Spec.	Acc.
CSF	Hybrid SVM	99.74	76.07	95.65		99.85	66.47	89.79
	ICA+SVM	99.77	53.42	87.53		99.82	48.11	82.43
GM	Hybrid SVM	87.50	98.41	89.97		94.35	72.21	88.65
	ICA+SVM	89.95	78.29	82.91		96.44	48.68	78.23
WM	Hybrid SVM	99.74	99.35	99.69		90.64	84.88	89.95
	ICA+SVM	97.87	99.31	98.82		95.70	66.07	75.74
WML	Hybrid SVM					99.93	99.98	99.95
	ICA+SVM					99.94	52.1	86.47

Sens. = Sensitivity, Spec. =Specificity, Acc. =Accuracy

From classification results, an overall performance improvement observed for hybrid approach compared to ICA based classification. In general, both the classifiers show high sensitivity, but specificity is found to be less. ICA based SVM showed performance measures different from results in Chapter 3 and [29], since it is based on IC values corresponding to each brain tissue sample. For CSF, accuracy improvement by the new method was +8.12 in

normal case, whereas abnormal case data showed an increase of 7.36 in accuracy. Presence of MS lesions decreases the performance of all tissues from abnormal data. In the case of GM, sensitivity is slightly higher for ICA+SVM, but specificity is too low compared to hybrid SVM. Normal WM tissue classification methods by both methods presented almost the same performance. However, in the case of abnormal data, hybrid SVM exceeded ICA+SVM in accuracy by a value of 14.21, but showing a dip of 5.06 % in sensitivity. An excellent performance was observed for hybrid SVM in the case of WML classification, with very good values for all measures. ICA+SVM could give comparable sensitivity values, but showed a great difference in specificity, only 52.1%.

Synthetic images are from simulated database, and are generated with almost uniform data characteristics for different slices. So the evaluation using these dataset can be considered as only a preliminary study. A detailed evaluation of the clinical data can give a good picture of the performance improvement in brain tissue classification by hybrid SVM.

6.4.2 Clinical image analysis

In the case of normal data, 360 multispectral sets were formed from 120 training cases. The 48 cases from test dataset were considered for performance evaluation. Three multispectral sets were considered for each case and the results were averaged. Considerable variations observed in results from patient to patient. To analyze the complete behavior of the classification, standard deviation (std.) of values corresponding to each performance was also considered. Small values for standard deviations indicate less variability. That means the consistency or stability of the algorithm in the classification increases with decrease in standard deviation. Average and standard deviation of these 48 cases are summarized in Table 6.3. Same classification process was repeated for samples selected from

conventional ICs in the same environment, and results are tabulated in Table 6.3.

Out of 110 WML cases, 75 were considered for training phase with 225 multispectral sets formed from it. Rests of the 35 cases were considered in the test phase, and classification was conducted as described for normal case. We divided other abnormal cases into a group of 75 cases for training, and a group of 29 cases for testing. Average and standard deviation of the performance measures for these 35+29 cases are added in Table 6.3.

Table 6.3 Clinical image analysis

		Normal				Abnormal		
Tissues	Method	Sens. ±std.	Spec. ±std.	Acc. ±std.		Sens. ±std.	Spec. ±std.	Acc. ±std.
CSF	Hybrid SVM	98.42± 0.65	84.08 ±11.45	95.75 ±2.10		97.23 ±2.17	51.15 ±13.05	85.11 ±3.54
	ICA+SVM	91.78 ±3.97	43.41 ±12.70	81.58 ±3.06		90.19 ±6.47	43.31 ±2.06	66.55 ±7.13
GM	Hybrid SVM	87.30 ±12.36	92.10 ±4.52	90.75 ±4.95		98.38 ±0.59	52.90 ±4.79	91.13 ±2.04
	ICA+SVM	85.94 ±12.93	74.65 ±5.11	83.93 ±10.04		96.24 ±2.57	39.54 ±7.43	83.54 ±2.89
WM	Hybrid SVM	91.56 ±7.48	89.29 ±02.47	92.21 ±2.24		94.92 ±0.18	73.60 ±11.78	90.94 ±1.60
	ICA+SVM	89.67 ±9.95	68.94 ±23.76	81.36 ±7.69		98.35 ±0.6	50.32 ±6.25	74.82 ±7.22
Abnormal tissues	Hybrid SVM					96.12 ±1.80	95.45 ±1.07	95.58 ±1.5
	ICA+SVM					98.15 ±0.49	47.17 ±5.27	71.24± 9.07

Sens. = Sensitivity, Spec. =Specificity, Acc. =Accuracy, std. = standard deviation

ICA+SVM classification was also performed in the same environment for result comparison, and measured values are included in Table 6.3. Proposed method could present comparatively good results for all cases. An overall

performance decrement is observed in the case of brain tissue classification from abnormal data, in comparison with those from normal database.

- Considering CSF, sensitivity and accuracy values from proposed method were found to be more consistent for normal and abnormal case with good improvement, but specificity found to be varying on different real data. In the case of normal CSF, hybrid SVM showed an improved performance compared to ICA+SVM, showing average sensitivity/accuracy of 98.42/95.75 against 91.78/81.58, with less standard deviations. Almost similar behavior observed in the case of abnormal CSF also, but specificity was found to decrease to around 51%, with increased inconsistency.

- In the case of GM classification using hybrid SVM, approximately +8% accuracy improvement over ICA+SVM was observed for normal and abnormal data. Standard deviation values for hybrid SVM indicated the improvement in consistency also.

- ICA+SVM could yield comparatively good average sensitivity for abnormal WM classification, showing an average 98.35% against 94.92% from hybrid SVM. However, the specificity in presence of abnormalities could not give such a high value. Hybrid SVM was observed with a consistent high accuracy, showing an average of 90.94 against 74.82.

- Abnormal case sensitivity values are more promising, but specificity got reduced due to presence of abnormal tissues, especially in the case of hyper intense lesions. The average sensitivity was observed to be slightly higher in the case of ICA+SVM. But the poor value observed for specificity, 57.17, discourages ICA+SVM as an accurate classifier in brain tissue analysis. Hybrid SVM yielded very good values for sensitivity/specificity/accuracy, 96.12/95.45/95.58, with standard

deviation always less than 2. Improvement in abnormal tissue analysis was found to be most attractive with a drastic increase of accuracy value from 71.24% to 95.58%, with improved consistency confirmed by reduced standard deviation from 9.07 to 1.5.

A separate performance evaluation using ROC curves [77] for WML and other abnormalities are shown in Fig. 6.4 and Fig. 6.5. Fig. 6.4 is the performance curve (TPR vs. FPR) using 1740 samples selected from 29 abnormal test cases excluding WML. Area Under the Curve (AUC) for hybrid SVM was measured as 0.9580, whereas ICA+SVM provided a poor performance with low AUC value 0.7228.

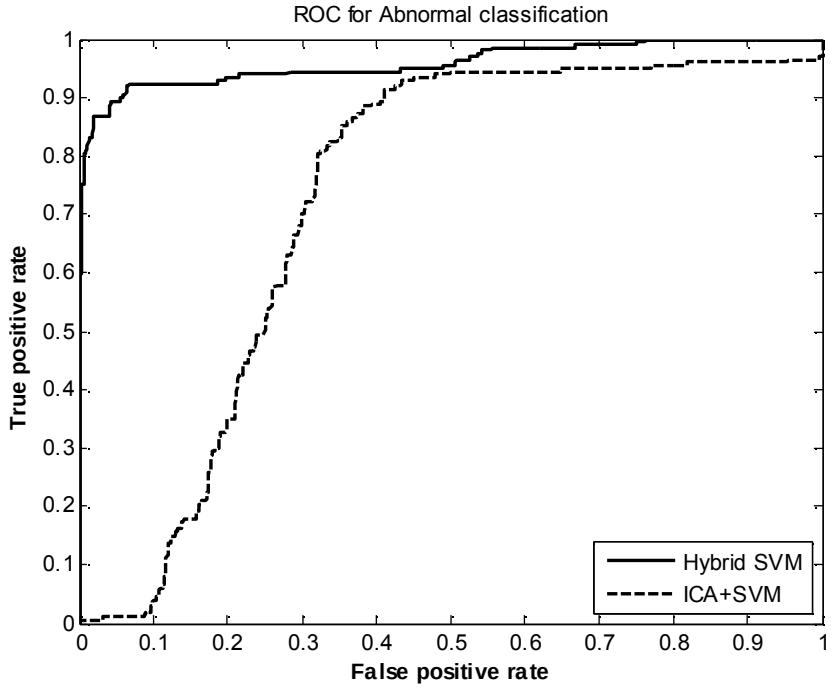


Fig. 6.4 ROC curve for abnormal cases by hybrid SVM and ICA+SVM

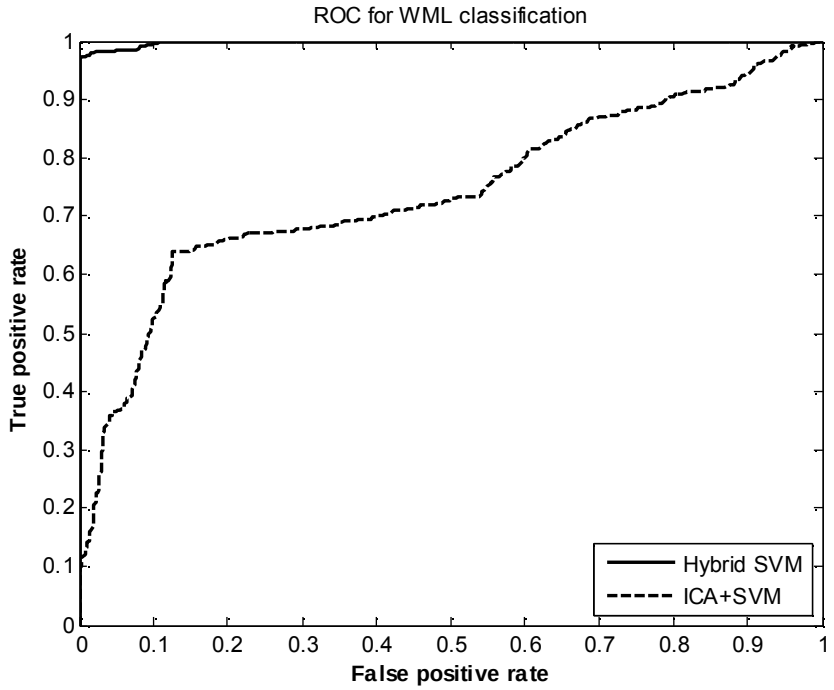


Fig. 6.5 ROC curve for WML by hybrid SVM and ICA+SVM

Fig 6.5 plots the performance curve (TPR vs. FPR) using 2100 samples selected from 35 WML test cases. AUC for hybrid SVM was measured as 0.9982, whereas ICA+SVM provided a poor performance with low AUC value 0.7403. In both cases, AUC values close to 1 from ROC curves give evidence for the high performance of hybrid SVM classifier for abnormality analysis.

6.4.3 Performance analysis with daubechies wavelets

Chapter 4 discussed that different wavelets provide varying performance results in brain tissue classification. Wavelets from daubechies family were considered in this section, to study the effect of the smoothness variation by vanishing moments [163] in quantitative analysis of brain tissue classification.

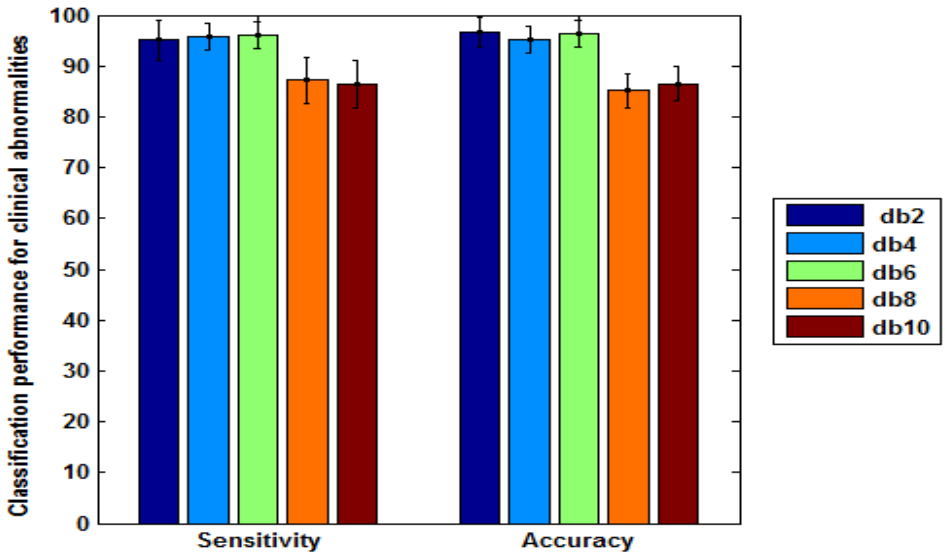


Fig 6.6 Error bar diagram showing classifier performance variations for different orders of daubechies wavelets

In this experiment, same data and procedure as described in section 6.4.2 were followed in feature extraction. Samples were categorized into a binary group, abnormal tissues and normal tissues. Total 64 cases were considered in the testing phase. Average and standard deviation of the measured sensitivity and accuracy are shown as an error bar diagram in Fig. 6.6 for different orders 2, 4, 6, 8, and 10. Low order wavelets are found to be yielding best performance with more consistent values for sensitivity and accuracy. Db6 could improve both sensitivity and accuracy to greater than 95% with less standard deviation. Db8 and Db10 results were found to be varying in the range [80% - 90%].

6.5 Qualitative analysis of brain tissue segmentation

In clinical trials, quality and accuracy of results from original sequences for visual classification is very important. Location and shape of the abnormalities as well as the effect of these abnormalities in other normal brain portion has great significance in clinical diagnosis. In this section we

visually analyze the classified tissues from hybrid SVM and ICA+SVM with the help of synthetic and clinical data. Classification models selected in section 6.4 were utilized to classify each sample in the multispectral set. GM is avoided because of the vagueness in samples from input images. Input image itself provides uncertainties for some tissue samples, which cannot be determined by an experienced radiologist or by a computer aided method. Classification process often neglects this fuzziness in data, which will reduce the efficiency of the method. Visual assessment of the improvement in quality is explained in this section with abnormal case data.

6.5.1 Segmentation of synthetic brain MRI

Axial slices containing presence of MS Lesions (WML) are shown in Fig 6.7. T1WI shows WM details, with patches of MS lesions. T2WI shows CSF, MS lesions, GM and other extrameningial tissues. PDI is used to locate the lesions with more information. ICA and multisignal wavelet analysis was performed on these input datasets for feature extraction. ICs from conventional ICA are shown in Fig. 6.8. IC1 gives details of CSF and WML; GM details can be extracted from IC2, and IC3 shows WM details. Minimum and maximum value represented by each IC is also given in Fig. 6.8.

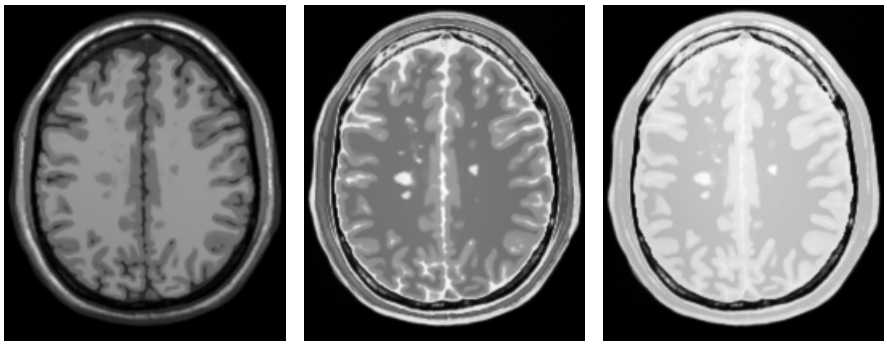
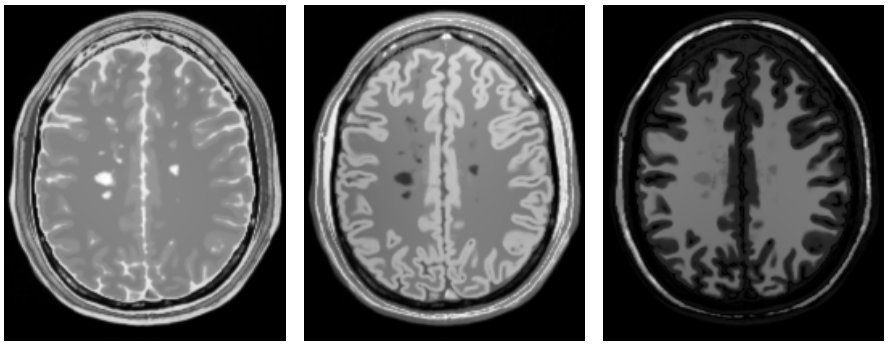


Fig .6.7 Synthetic input slices; T1WI, T2WI, and PDI from left to right

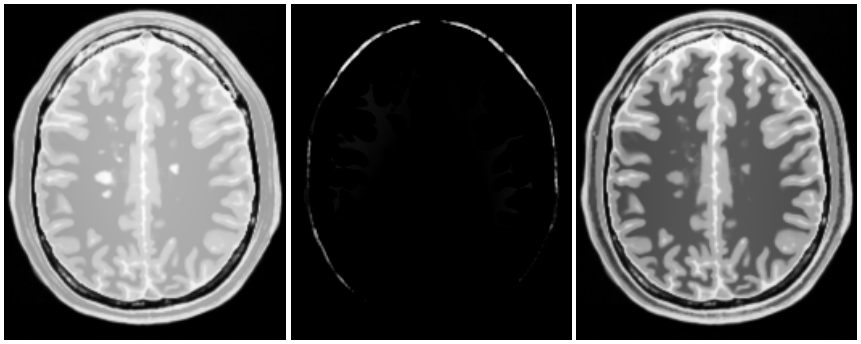


IC1
Min = 0
Max = 763.5293

IC2
Min = 0
Max = 656.6606

IC3
Min = 0
Max = 659.2381

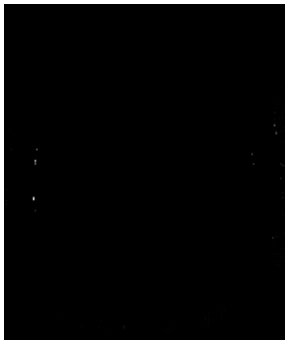
Fig.6.8 Independent Components (IC1, IC2, IC3), with range of values



Approx. coefficient
Min= 0
Max= 1.9250e+005

Detail coefficient 1
Min= -1.2693e+004
Max= 4.0962e+003

Detail coefficient 2
Min= -185.4580
Max= 2.3325e+004



Detail coefficient 3
Min= -1.3115e+004
Max= 182.5386

Fig .6.9 Approximation (AC) and detail coefficients (DC1, DC2, DC3)
from multisignal wavelet analysis

Selected multisignal wavelet coefficients according to the algorithm 6.1 from multisignal wavelet analysis using db2 wavelet are shown in Fig 6.9. From the observed range of values for each coefficient it is obvious that each coefficient can provide a wide range of values for different samples. Approximation Coefficient (AC) gives details of WML, where Detail Coefficient 1 (DC1) shows WM details with minimum -1.2693e+004 and maximum 4.0962e+003. DC2 is found to be good for features of CSF and GM. For DC3 some unique values were provided, but it is not visible through the pixel range for display.

Classified tissues from proposed classification and ICA+SVM are shown in Fig. 6.10. First row represents CSF, second row shows GM, third row gives WM, and last row shows the lesion details. Groundtruth (GT) images are provided in the first column as a reference to analyze the observed results. All these results were presented after removing extrameningial tissues to perform a better comparative study. IC1 and DC2 were selected to generate CSF. IC2 and DC2 gave GM; IC3 and DC1 could classify WM, and IC1 and AC generated the white matter lesions (WML).

On analyzing the first row, it is very clear that classified CSF from proposed method is more similar to GT than CSF from ICA+SVM. However, both methods show less details compared to GT. Traces of WML are also shown in results from ICA+SVM. Considering GM, more negative pixels are considered as positive by both the methods, but comparatively better tissues were obtained by the proposed method. It is very clear from 3rd and 4th row that WML and its effect on WM are better described by proposed method in comparison with ICA+SVM. WM tissues from proposed method identify more number of lesion locations (3rd row, black pixels in WM) compared to those from ICA+SVM. But, information loss is observed for WM from proposed method (circled portion in Fig. 6.10, 3rd row, 2nd column).

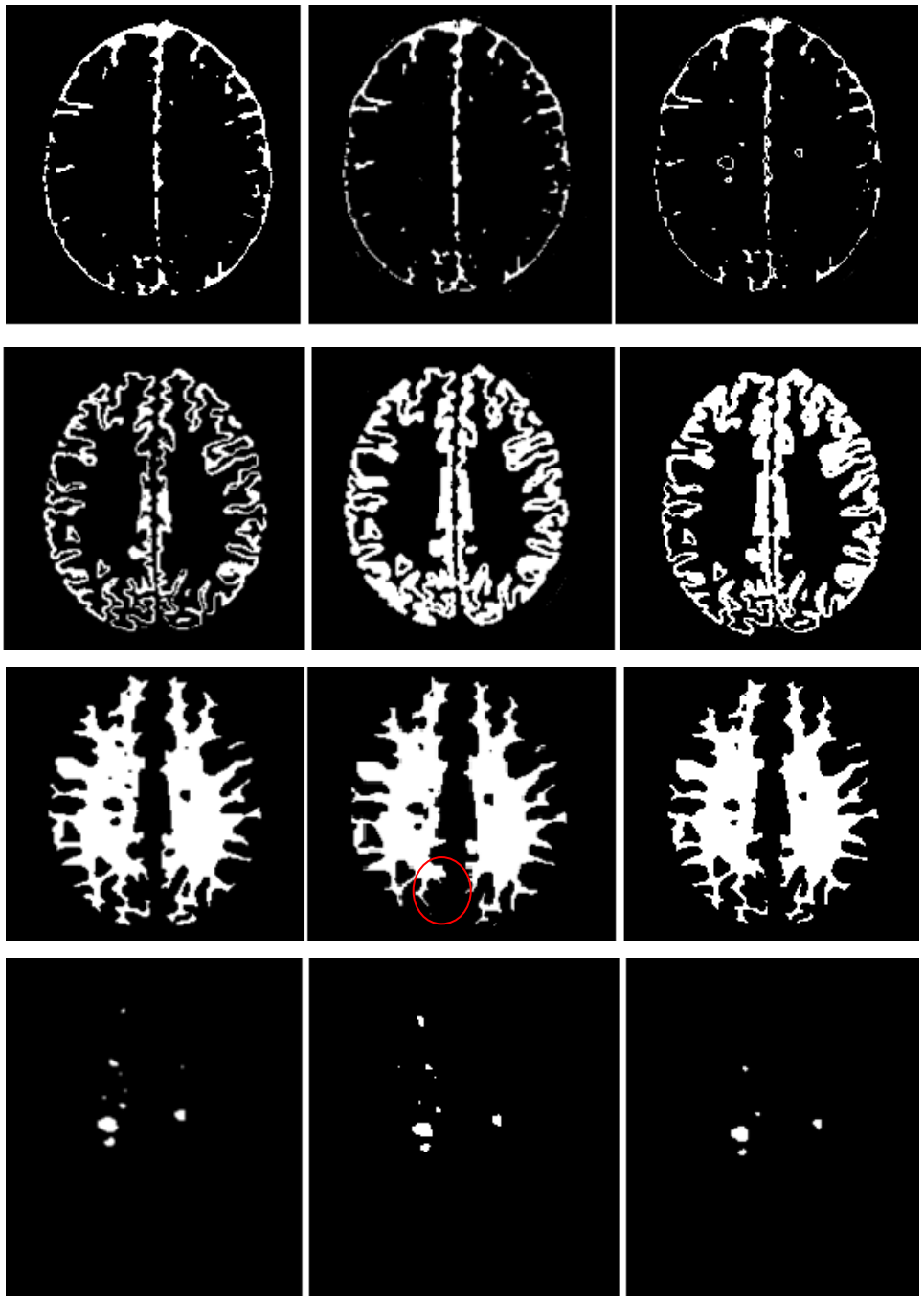
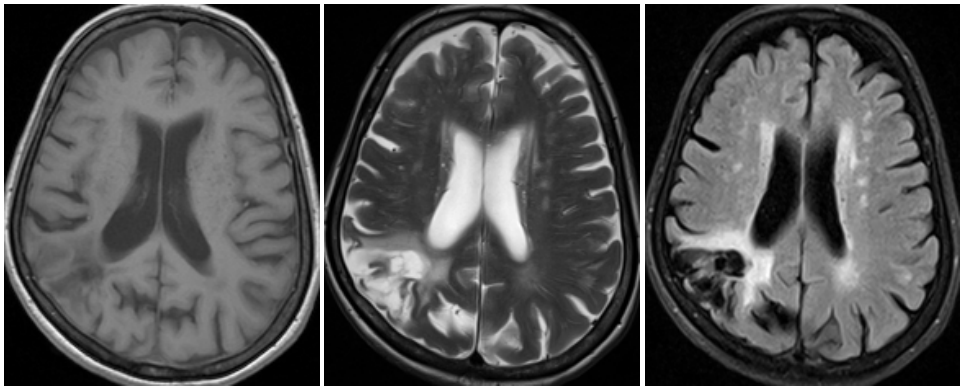


Fig 6.10 Classified tissues, CSF (1st row), GM (2nd row), WM(3rd row), and WML(last row) in the order of GT, results from proposed classification and ICA from left to right.

Available number of lesions and amount of information by proposed method is found to be significantly greater than that from ICA+SVM (last row). ICA+SVM could identify the major lesions, but it neglects all the small lesions.

6.5.2 Qualitative analysis of clinical dataset

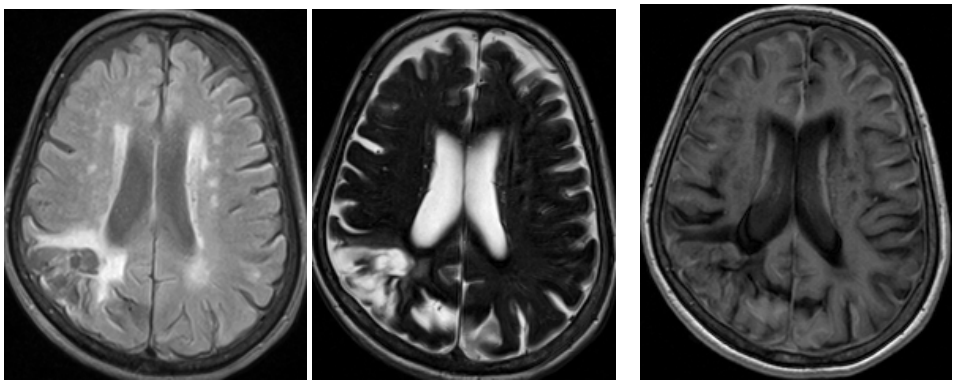
In the qualitative analysis of clinical data, images showing evidence of chronic infarct with gliosis noted involving right parieto-occipital region of brain were selected. Parenchymal signal in the area of chronic infarction continues to show CSF-like hypo intensity on T1WI and hyper intensity on T2WI as shown in Fig. 6.11. Multiple small chronic lacunar infarcts are also noted in deep white matter of both cerebral hemispheres (better noted in FLAIR image). The core of the chronic infarction is also CSF-like on FLAIR images, but surrounding gliosis appears hyper intense. Lacunar infarcts, also known as lacunar strokes, are small areas of dead tissue found deep within the brain.



T1WI T2WI FLAIR

Fig 6.11 Clinical inputs

When we analyze the tissues, T1WI shows WM. But presence of abnormalities is not reflected in the image. In T2WI, we cannot distinguish between CSF and abnormalities.

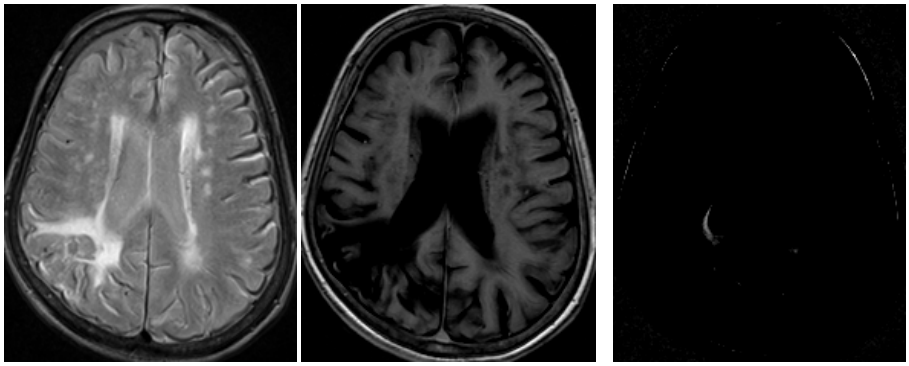


IC1
Min = 0.7562
Max = 822.34

IC2
Min = 0.0032
Max = 558.8671

IC3
Min = 1.6267e-004
Max = 956.74

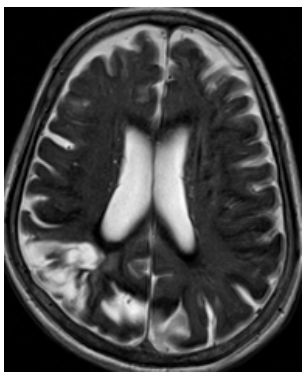
Fig. 6.12 Clinical Independent Components



AC
Min= -73.6715
Max= 1.1729e+003

DC 1
Min= -88.9931
Max= 201.9797

DC 2
Min= -396.6071
Max= 75.0493



DC 3
Min= -65.7962
Max= 331.1719

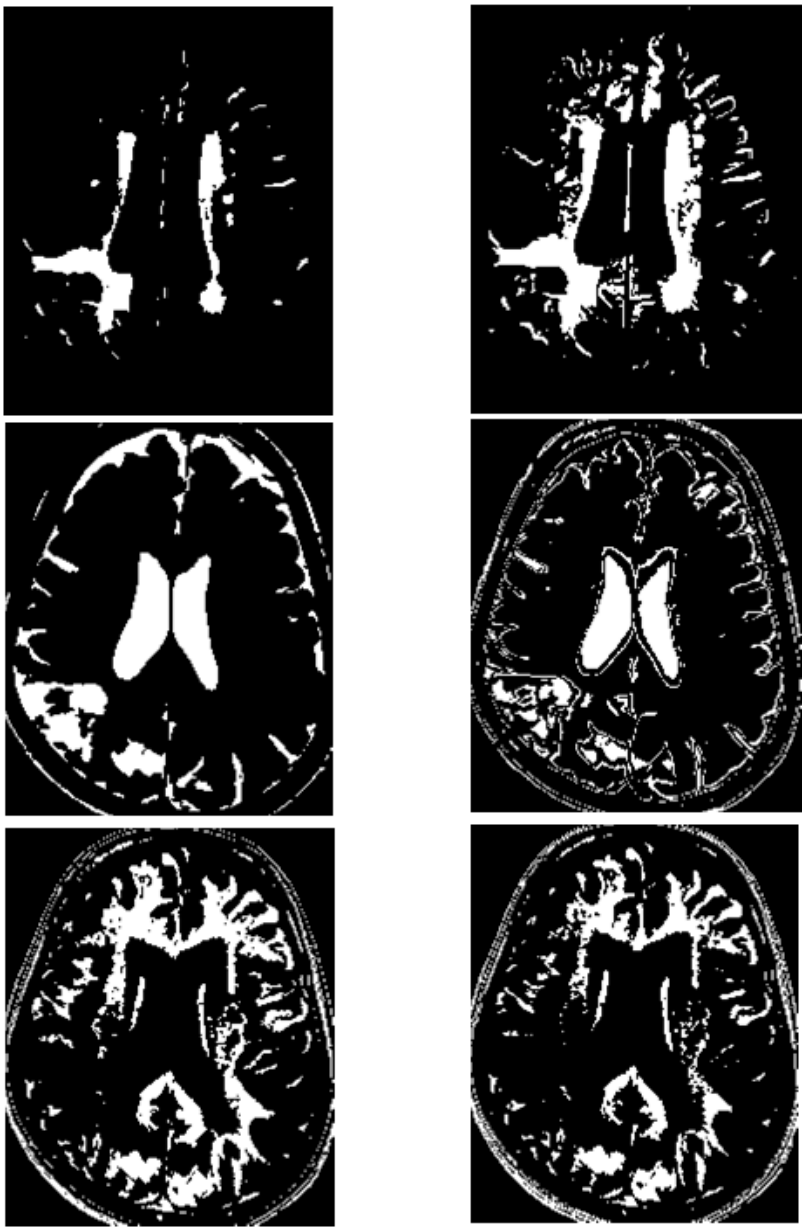
Fig. 6.13 Coefficients by multisignal wavelet analysis

Lacunar infarcts found in FLAIR sequences are too small and mixed with its background, which makes feature extraction very difficult. Not much detail on GM was available from these sequences. A simultaneous analysis of these images is a challenge in conventional applications.

Multisignal wavelet analysis using *db6* and ICA were applied on the co-registered images in Fig. 6.11. ICA results are shown in Fig. 6.12 and Wavelet coefficients are presented in Fig 6.13. Small chronic lacunar infarcts and surrounding tissues of chronic infarction, which are found to be hyper intense in FLAIR images are available from IC1. IC2 extracts CSF and CSF-like abnormalities. IC3 shows WM details, but abnormal tissues are not found to be showing their presence in WM portion.

AC coefficient is found to be good in providing information on small and large lesions affecting WM. DC1 gives a complete picture of WM, with all of the affected area got removed. Information inherent in DC2 is not visible because of the negative values. DC3 provide CSF details and CSF like abnormalities (chronic Infarctions). Range of coefficient values gives the evidence of uniqueness in representing different basic tissues.

Classification results by hybrid SVM and ICA+SVM are summarized in Fig. 6.14. Extrameningial tissues are not removed from the classified results. No gold standard was available as groundtruth in the case of clinical case, since a certain extent of fuzziness exists for each tissue. Manual segmentation results from experienced radiologists were often considered as the reference for comparison. Top row shows the classified abnormalities, small chronic infarcts in the deep white matter and surrounding hyper intense lesions. AC coefficient and IC1 were included in the classification of these tissues. It is observed that hybrid SVM results provide a better representation of abnormalities compared to ICA based SVM results.



Hybrid SVM ICA+SVM

Fig. 6.14 Clinical brain tissue analysis
First row- White matter lesions, Chronic lacunar infarcts,

Second row- CSF and chronic infarction

Third row- White Matter

ICA based SVM shows uncertainties in prediction, which caused the misclassification as observed in the result. Considering the second row, we see that the CSF and CSF-like abnormalities are clearly segmented by hybrid SVM. ICA based SVM failed to generate the positive tissues (here, CSF or CSF like abnormal tissues), but more negative tissues are found to be misclassified as positives in the result. IC2 and DC3 participated in the prediction of these tissues. From FLAIR and T2WI in Fig 6.11 it is very clear that a large portion of the WM is affected by the described abnormalities. When we consider the abnormalities in WM and remaining normal WM (last row), hybrid SVM yielded a good result with all the abnormal locations properly identified.

6.6 Summary

The newly proposed hybrid feature selection method for improved brain tissue classification using SVM is discussed in this chapter. Multisignal wavelet analysis in the spectral domain is proposed to preserve the global and local information, with equal priority in feature extraction for multispectral MRI analysis. Tissue samples, represented by ICA and wavelet features, were selected for training and classification using SVM. Performance of the new algorithm was analyzed and confirmed by a detailed quantitative and qualitative analysis with ICA based classification.

The SVM models obtained from training and validation were used to segment the synthetic and clinical data. To ensure the effect of the method in identifying abnormalities and their effect on other tissues, a detailed visual analysis was also included in this chapter. Comparison with classification results from ICA demonstrated the performance improvement by the new method in multispectral brain tissue analysis.

Chapter - 7

COMPARATIVE ANALYSIS AND DISCUSSION

7.1 Introduction

In this chapter, we compare the classification results from proposed methods with conventional classification approaches, and summarize the results. From the previous chapters and literature, it was observed that supervised classification provides better performance in brain MRI analysis compared to FCM based segmentation. Therefore, we considered SVM based on proposed methods and conventional ICA in the detailed comparative study described in this chapter. In addition to these, state of the art algorithms, SVM, ANN, PNN, and Bagging were also included in the comparative study. Section 7.3 gives a brief discussion on the proposed methods and classifications, summarizing the merits and limitations in brain tissue analysis.

7.2 Comparison with existing approaches

In this section we focus on a more elaborative study using conventional classification methods. Detailed comparative analysis of nine supervised classification algorithms including proposed methods is presented in this section.

7.2.1 Experimental setup

Total 214 abnormal clinical cases were considered to conduct this experiment. 214x3 multispectral slice sets as described in Table 5.1 were considered to collect the features of CSF, WM, and abnormality classes. In the case of classifications based on feature extraction methods, 12840 feature

vectors were collected from unmixed ICs. For other conventional algorithms, these feature vectors were directly selected from the multispectral slice sets. 9000 out of 12840 were considered for training purpose, and rest of the feature vectors were utilized to generate the results for comparison. A 3x3 pixel window was used to locate the best features of a particular tissue as described in Chapter 5. For the implementation of proposed methods, we selected the optimal parameters observed from result analysis in Chapter 5. i.e., for SC-ICA we selected the threshold value 0.06, and daubechies type '*db6*' was selected for wavelet based analysis. For other conventional algorithms, optimal parameters yielding best results were selected by trial and error. Matlab implementations for 4-layer feed-forward back propagation network with three hidden layer sizes [10 8 5], probabilistic neural network with spread value 0.9, non-linear SVM with RBF kernel and default parameter settings, and ensemble method with 100 bagged decision trees were applied on training and testing feature vectors to complete the comparative analysis.

7.2.2 Comparative analysis

Average sensitivity, specificity, accuracy and standard deviation (std.) of these measures were observed with tumor and WM classes, and summarized in Table 7.1. Hybrid SVM, MW-ICA+SVM, SC-ICA+SVM, MICA+SVM and Bagging provided relatively high performance in classification. The main focus of this dissertation is to improve the pixel based brain tissue classification through efficient feature extraction methods, which highlights the presence of small abnormalities along with the global features. Results in Table 7.1 demonstrated this goal through results from proposed classification techniques.

Table 7.1 Performance comparison of supervised classification methods for clinical abnormal data

	Tumour/Lesion				White Matter		
Methods	Sens. ± std. (%)	Spec. ± std. (%)	Acc. ± std. (%)		Sens. ± std. (%)	Spec. ± std. (%)	Acc. ± std. (%)
SC-ICA +SVM	**94.70 ± 3.71**	**96.62 ± 2.51**	**95.05 ± 3.42**		**93.63 ± 3.71**	**92.54 ± 3.92**	**92.62 ± 3.35**
MICA+SVM	94.45 ± 4.42	96.71 ± 2.14	95.21 ± 3.52		83.74 ± 6.11	90.73 ± 3.44	89.42 ± 4.13
MW-ICA+SVM	**94.29 ± 2.91**	**97.47 ± 1.87**	**96.34 ± 1.82**		**86.38 ± 6.93**	**95.43 ± 1.35**	**94.91 ± 1.98**
Hybrid SVM	**96.24 ± 1.64**	**96.55 ± 1.37**	**96.49 ± 1.6**		**95.63 ± 1.23**	**93.60 ± 4.99**	**94.34 ± 2.42**
ICA+SVM	92.23 ± 3.98	84.42 ± 1.67	89.64 ± 2.87		81.84 ± 3.66	90.43 ± 4.96	88.53 ± 3.98
SVM	92.44 ± 4.53	61.11 ±13.43	75.12 ± 9.61		80.12 ± 5.32	70.46 ± 2.57	79.71 ± 1.94
ANN	88.68 ± 2.94	70.24 ± 1.39	83.61 ± 3.13		84.63 ± 4.37	81.21 ± 3.25	82.72 ± 3.55
PNN	94.63 ± 2.49	63.16 ± 3.44	83.11 ± 2.72		65.34 ± 3.23	92.07 ± 1.8	74.54 ± 3.01
Bagging	93.08 ± 3.95	84.64 ± 3.78	91.03 ± 4.12		91.72 ± 3.66	88.33 ± 3.71	90.21 ± 2.4

Sens. = Sensitivity, Spec. = Specificity, Acc. = Accuracy, std. = standard deviation

In the case of abnormality analysis, SVM based on proposed methods showed significant positive difference. Hybrid SVM provided high performance classification with best sensitivity/specificity/accuracy, 96.24/96.55/96.49, with corresponding standard deviations, 1.64/1.37/1.6, the least among all the methods under consideration. The second best classification was provided by MW-ICA+SVM, with almost similar

performance as hybrid SVM. MICA+SVM and SC-ICA+SVM also showed better performance compared to conventional approaches, with high accuracies, 95.21 and 95.05, respectively. Conventional Bagging approach was also found to be providing high accuracy, 91.03, but with more variation indicated by the value of standard deviation, ±4.12.

Regarding WM classification, overall performance of hybrid SVM and MW-ICA+SVM was observed as more promising. Hybrid SVM provided the best sensitivity with least standard deviation, 95.63 (± 1.23), whereas MW-ICA+SVM gave best specificity, 95.43 (± 1.35). Accuracy values (with standard deviations) demonstrated the efficiency of both algorithms yielding, 94.34 (± 2.42), for former method, and 94.91 (± 1.98) for later approach. MICA+SVM performed better than conventional methods except bagging, but failed to provide a high performance classification as it showed in the case of abnormal tissues. SC-ICA_0.06 also showed good accuracy, 92.62, with a standard deviation of ±3.35. In the case of WM also, bagging provided an accuracy >90% with less variations.

An error bar diagram can better depict these variations to evaluate the improvement in efficiency and consistency. Fig. 7.1 illustrates the general behavior of the accuracy results given in Table 7.1 with the help of an error bar chart.

From Fig. 7.1, performance improvement by proposed classifications is very evident without any indexing. SVM showed least accuracy with large inconsistency for abnormality analysis, whereas PNN showed least performance for WM classification. Even though experimental results recommended the potential of the new algorithms in brain tissue classification, some exceptional results also observed due to some specific and general limitations, which will be discussed in the next section.

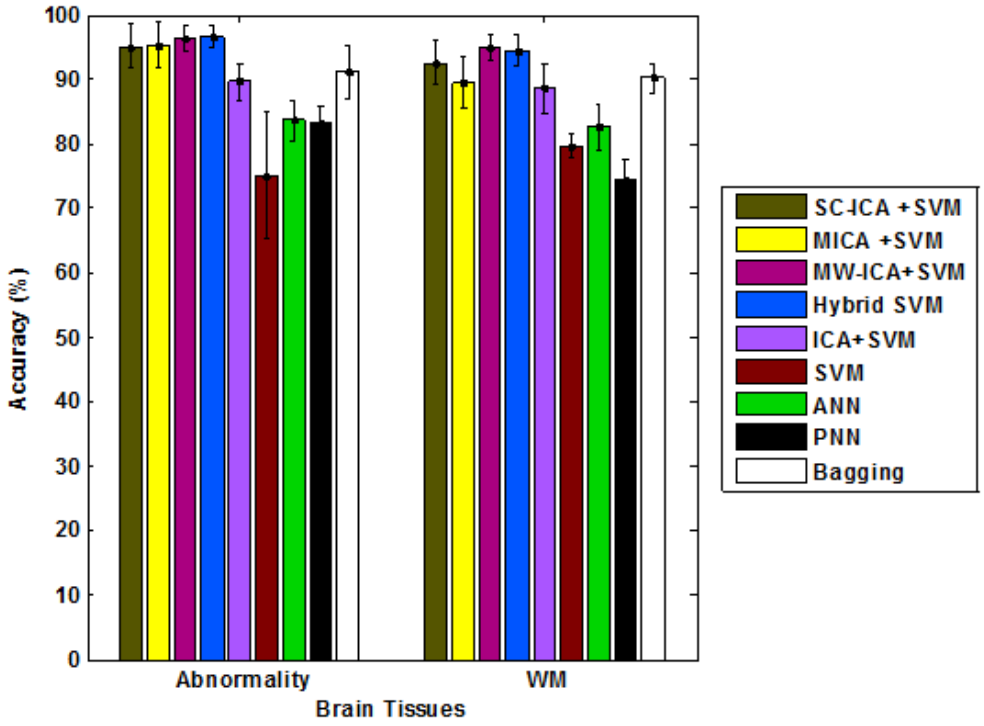


Fig. 7.1 Error bar diagram for performance comparison of supervised classification methods

7.3 Discussion

Both supervised and unsupervised classification approaches are found to be widely used in multispectral brain MRI analysis [7]. No operator intervention is required in the case of unsupervised analysis, and input multispectral cube is automatically segmented into different clusters. Sometimes this blind clustering without prior knowledge failed to produce meaningful segmentations, especially in the case of MRI analysis where many unknown brain tissue clusters with different tissue characteristics may be present. However, they are very successful in clinical applications of normal brain analysis, since it includes only the known structural characteristics [167]. Abnormalities with complex knowledge need some additional information from experts to provide a more accurate segmentation [168]. There we can

exploit the advantage of supervised methods to achieve the superior performance. The main issue with the supervised classification is the inconsistency due to large variations in intra-operator and inter-operator feature measurements [29]. However, high performance feature analysis techniques can significantly reduce the inconsistent results. Experimental results in Chapter 5 demonstrated the efficiency and accuracy of SVM over FCM in MRI analysis.

Compared to other conventional classification methods such as Gaussian maximum likelihood classifier and neural networks, SVM shows high generalization capability with relatively small number of training samples [29]. However, selection of non-linear kernels and optimal parameters highly influences the classification performance [5]. Error bar diagram in Fig. 7.1 implies the necessity of some pre-processing methods like ICA to achieve a high performance classification. Improvement in accuracy by feature selection through ICs, for abnormalities and WM is very evident in Table 7.1 and Fig 7.1. As a global transform, ICA finds its difficulty in extraction of less frequently occurred information from massive amount of data [53, 152]. This dissertation proposed a few methods based on spectral clustering and multisignal wavelet analysis, as discussed from Chapter 4 to Chapter 6. Detailed quantitative and qualitative analysis was also conducted to confirm the potential of the new methods in supervised and unsupervised brain tissue analysis. The following section summarizes the merits and demerits observed on the basis of experimental analysis.

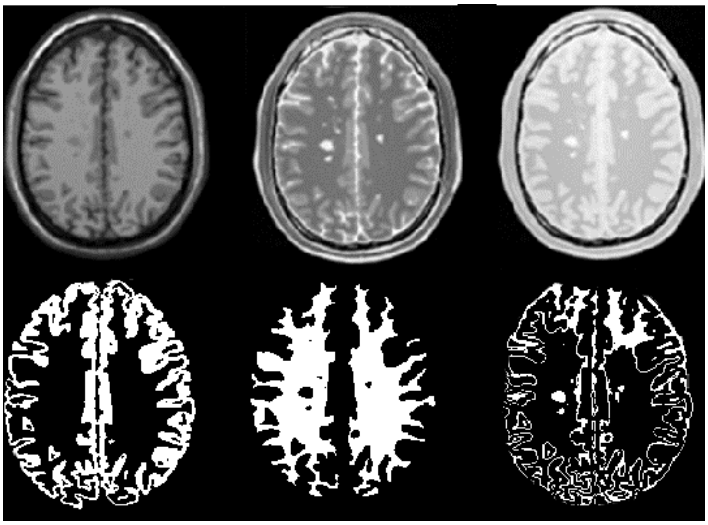
7.3.1 Benefits of SC-ICA based classification in MRI analysis

In proposed SC-ICA, a clustering algorithm based on spectral angle distance is used to perform object based ICA on multispectral data. The proposed algorithm was validated with synthetic and real MRI data. Compared to ICA based SVM for MRI analysis [29], SC-ICA+SVM provides better

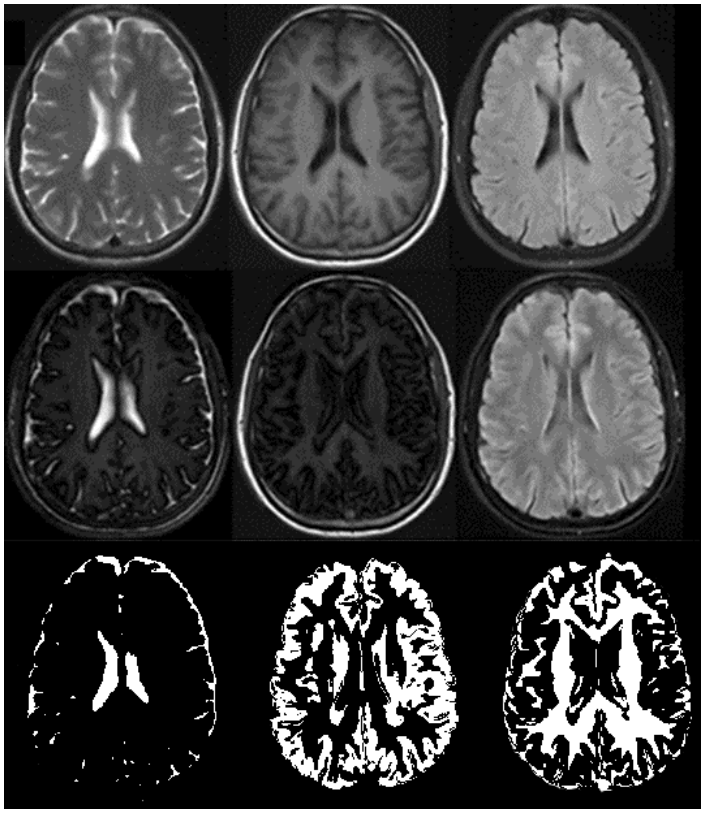
performance in small object detection. In abnormal clinical analysis, SC-ICA based classification outperforms ICA+SVM with best results (Fig. 5.16 and Fig. 5.18) and improved TI values (Table 5.5 and Table 5.7). It is demonstrated by observed sensitivity/TI values, 90.3%/0.75 for SC-ICA_.03+SVM, against 84.9%/0.52 for ICA+SVM (Table 5.6) in lesion analysis. The positive impact of SC-ICA+SVM in locating presence of abnormalities on other tissues is also demonstrated through highest TI value for WM, 0.89 and 0.90 (Table 5.5 and Table 5.7), in supervised and unsupervised clinical image analysis. High accuracy values observed for SC-ICA_.03+SVM and SC-ICA_.06+SVM (Table 5.6), (92.57±6.25) % and (82.31±14.9) %, are found to be very promising in clinical abnormal tissue analysis. From Table 7.1, it is observed that comparatively less standard deviations and better average sensitivity/specificity/accuracy values (%), 94.7/96.62/95.05, were shown by SC-ICA+SVM. These results indicate the stability and superiority of the method over conventional approaches in brain tissue classification.

7.3.2 Limitations of SC-ICA based classifications

Threshold value selection plays a significant role in accuracy of classified results from SC-ICA. Experimental results in Chapter 5 demonstrated that classification performance highly varies on selected threshold values. Low threshold values can improve classification of local features. But it may lead to over-clustering, which adversely affects normal tissue analysis. Generated IC and distorted WM tissue (Fig. 7.2(a)) from SC-ICA_.03+SVM explains this with the lowest TI value, 0.51. However, improved sensitivity value, 97.35%, in Table 5.3 suggested that lesion details were not lost from the result; low TI value was generated as a result of misclassification due to over clustering by SC-ICA_.03.



(a) SC-ICA_.03_SVM on Synthetic images



(b) SC-ICA_.03_SVM on clinical normal case

Fig. 7.2 Low threshold (0.03) effects from SC-ICA

Input data characteristics have a big role in optimal threshold selection. In general, SC-ICA_.06 based classifications performed well for brain tissue analysis. However, specific analysis on normal and abnormal clinical database revealed a different picture. SC-ICA_.1 based classification provided good results for normal tissue analysis, but optimal values were observed from SC-ICA_.12 for abnormal tissues. Normal case results for SC-ICA_.03 was not satisfactory in visual and quantitative analysis (Fig 7.2(b)). However, it was highly attractive for abnormality analysis, which was emphasized through sensitivity, specificity and accuracy values of abnormal tissues, as given in Table 5.5 and 5.6.

Feature extraction cost due to clustering in SC-ICA is another issue. Table 4.3 stated that time complexity is directly proportional to number of clusters. Decrease in threshold will increase the number of clusters. For example, a threshold value 0.03 will generate more than 833 clusters from a clinical abnormal multispectral set. The required time in that case was observed as 780 seconds, whereas ICA took only 0.7 seconds for feature extraction. On increasing the threshold to 0.15, the execution time was found to be decreasing to 40 seconds (Table 4.3).

7.3.3 Benefits of wavelet based approaches in MRI analysis

Wavelet based ICA extensions are proposed as an alternate to spectral clustering to retain the local information in brain tissue analysis. Experimental results demonstrated that multiresolution analysis avoids the loss of significant MRI details in feature extraction, and it provides a relatively good pre-processing step for high performance tissue classification and abnormality analysis. Low order daubechies wavelets were observed as yielding best performance in MRI analysis. No serious issues like threshold selection in SC-ICA observed for multisignal wavelet analysis.

MICA based classifications were observed as very good for abnormality analysis with less computational overhead. Visual results from Fig. 5.16 and Fig. 5.18 demonstrated the superior quality of the classified abnormalities. It can be seen from Table 4.3 that feature extraction for these good quality results can be done within 5 seconds. However, its poor performance in normal tissue classification needs refinement with more sophisticated algorithms.

The new method, MW-ICA, is a successful approach in unmixing each brain tissue into separate ICs, from which a reliable supervised or unsupervised classification can be easily performed. It combines two concepts to improve the brain tissue analysis; multisignal wavelet analysis helped to retain the local and global features with the same priority, and over complete-ICA issue is solved through band expansion using the reconstructed images from detail coefficients. In comparison with ICs from widely used ICA method (Fig. 4.12(c)), MW-ICA provides more tissue specific components (Fig. 4.12(b)). Lesion detection from ICA results was found to be difficult, since CSF and lesion details were accumulated in first component (Fig. 4.12(c) 1st column). In addition to that, ICA results failed to locate the presence of lesions in WM (Fig. 4.12 (c) last column) because of background dominating effects from global features. The closest agreement between segmented brain volumes from the groundtruth and MW-ICA results, observed from Bland-Altman plots (Fig. 5.5 - Fig. 5.8), recommends MW-ICA based segmentation as the best among the proposed methods. Improvement in MW-ICA+FCM over other competent methods was confirmed by the reduced mean bias (which is closer to zero) and shortened limits of agreement.

Improved sensitivity and accuracy values observed for all tissues demonstrated the efficiency and robustness of the new algorithm in clinical trials (Table 5.6 and Table 7.1). Increased average TI value (Table 5.5 and

Table 5.7) and reduced misclassification rates (Fig. 5.14) supports the effectiveness of the method in small abnormality analysis. Considerable reduction in standard deviation of these measures (Table 5.6, Table 7.1 and Fig. 5.14) implies the consistency of MW-ICA based classification in clinical analysis.

Hybrid SVM defines a unique feature set for each sample, with ICA and wavelet features for validation and classification using SVM. Classification results in Fig 7.1 showed the superior performance of hybrid SVM with improved efficiency and consistency. Significant improvement in statistical analysis (Table 6.3), increased AUC from ROC curves (Fig 6.4 and Fig.6.5) and refined quality of the segmented tissues from classification models (Fig. 6.14) recommends Hybrid SVM as a promising and potential supervised approach in normal and abnormal tissue analysis.

7.3.4 Limitations of wavelet based ICA extensions

Modified MICA based classification is robust and reliable in abnormality analysis. Experimental results in Chapter 5 demonstrated that multiresolution analysis avoids the loss of significant MRI details in feature extraction, and it provides a relatively good pre-processing step for high performance tissue classification. However, it suppresses majority information before wavelet reconstruction as discussed in step2 of MICA algorithm [53]. This will adversely affect the classification performance of normal brain tissues in abnormal multispectral data. For example, poor classification performance was observed for WM and GM tissues in Fig 5.16 (c) and 5.18 (d). Results in Table 5.4 and Table 5.5 support this observation with low TI values, 0.81/0.71 and 0.78/0.81, for synthetic and clinical GM/WM classification respectively.

Performance of wavelet based approaches, MW-ICA and hybrid SVM, are found to be varying on different values of parameters like wavelet type, degree of decomposition, number of bands in input signals etc. Variation in quantitative measures, TI and error rate for different wavelets, indicated that wavelet selection has an important role in quality of the results. Fig. 5.9 recommends *db12* for high performance classification with MW-ICA, whereas Fig.6.6 indicates *db6* as the optimal wavelet type for hybrid SVM.

Considering the cost of feature extraction, MW-ICA was observed as more time consuming. We measured 0.7 seconds for ICA, 4.9 seconds for MICA and 7.7 seconds for MW-ICA with *db8* wavelet in a typical clinical imaging analysis, on a Windows7 PC with Pentium Dual CPU of 2.0GHz/2GB RAM (Table 4.3). It can be varied on different values of parameters like wavelet type, degree of decomposition, etc., which is to be explored in future works. Multisignal wavelet can be effectively performed with more number of bands in input multisignal. Real environment imaging issues such as gaussian noise, intensity inhomogenity issues, motion artifacts were not considered in this analysis.

7.4 Summary

An overall summary and analysis of the proposed methods and conventional classification methods are included in this chapter. A detailed comparative analysis of supervised classification based on proposed ICA extensions and hybrid feature selection is discussed at the onset of the chapter. Following that, a detailed discussion on merits and demerits of spectral based approach and wavelet based approaches are included.

Classifications based on hybrid SVM and MW-ICA presented superior performance among the proposed methods. SC-ICA with optimal threshold is also found to be yielding competitive classification results. MICA based

analysis showed its potential in small abnormality classification, but for some global tissues it failed to provide better performance over conventional approaches. Experimental results recommended the efficiency of multispectral MRI analysis based on proposed methods in retaining local and global characteristics to improve the brain tissue classification. Potential of these methods in simultaneous analysis of small abnormalities like WML, and their effect on other tissues can be best exploited in disease progress evaluation and treatment.

Future work focuses on an adaptive threshold selection scheme that can solve the threshold and time complexity issues for spectral clustering system. Effectiveness of wavelets to work with different types of noise and artifacts are also under consideration as future study. Clinical accuracy of results can be improved significantly by addition of high informative MRI sequences to input multispectral data.

Chapter - 8

CONCLUSIONS AND FUTURE WORKS

8.1 Conclusions

In this thesis work, we explored the issues in ICA for brain tissue classification. Chapter 1 provided the problem definition, motivation and background introduction. Chapter 2 explained the literature and state of the art algorithms in MRI analysis, and Chapter 3 examined the ICA issues in MRI analysis with a detailed explanation on its background. First, it was noted that ICA neglects local information while dealing with a massive amount of data. Another major issue was over-completeness problem, in which more than one component accumulated in the same independent component, making it difficult to extract the relevant features accurately. In Chapters 4–7, various aspects of different solutions were explained, with the help of a thorough analysis using synthetic and clinical MRI data. In addition to that, open questions and potential future perspectives were also identified.

Chapter 4 discussed the three new ICA extensions to improve the classification performance in brain tissue analysis; A new Spectral Clustering ICA (SC-ICA), a new Multisignal Wavelet ICA (MW-ICA) and modified Multiresolution ICA (MICA). Each method was explained with relevant theory, algorithm and examples. Computational overhead on varying parameters is also compared. To justify the improvement by new methods, a comparative study with ICA is also added for each experiment. Quantitative and qualitative analysis of these proposed ICA extensions with supervised and unsupervised classifications, SVM and FCM respectively, is illustrated in Chapter 5. A detailed description on the synthetic and clinical, normal and

abnormal datasets used, and preparation of raw data for feature extraction was also added.

Although the ICA extensions were found to be superior to ICA in MRI analysis, some limitations were also observed to enable further improvement. Main issue was the influence of optimal parameter selection in classification performance. A supervised method with optimal parameters and best training set was found to be more desirable in MRI analysis than an unsupervised approach. It was another open question, whether best features from ICA and multisignal wavelet transforms can be successfully combined to improve the supervised brain tissue analysis. Hybrid SVM discussed in Chapter 6 gave a solution to this question, which was demonstrated through experimental results from normal and abnormal MRI data.

Chapter 7 gives an overall summary and analysis of the proposed methods and conventional classification methods. It was started with a detailed comparative analysis of proposed supervised classifications with existing approaches. After that, merits and demerits of spectral based approach and wavelet based approaches were discussed in detail. The complete dissertation work can be concluded as follows.

- A new Spectral Clustering ICA (SC-ICA), a new Multisignal Wavelet ICA (MW-ICA) and modified Multiresolution ICA (MICA) were proposed as ICA extensions to improve the normal and abnormal brain tissue classification from multispectral brain MRI.
- A new hybrid feature selection method based on ICA and multisignal wavelet analysis was designed and analyzed with SVM for brain tissue classification.
- The proposed methods were evaluated quantitatively and qualitatively, using clinical and synthetic MRI data, for normal and

abnormal cases. A comparative analysis with ICA based SVM, SVM, ANN, PNN and bagging was also carried out.

- Hybrid SVM, and classifications based on MW-ICA and SC-ICA showed high performance classification for both normal and abnormal datasets. MICA was found to be yielding best results in the classification of small abnormalities.
- Two major issues in ICA based MRI analysis have been addressed; poor performance in analysis of small abnormalities, and unmixing issues in separating more classes from less number of input images (OC-ICA).
- Experimental results demonstrated the potential of the proposed methods in pathological analysis of brain MRI.

8.2 Future directions

- Experimental results indicated that SC-ICA based analysis provide a high performance brain tissue classification, especially in the case of small lesions and tumors in clinical analysis. However, both classification performance and computational cost highly varies on selected threshold values. As a future work, a refined system using adaptive threshold selection scheme can solve the threshold and time complexity issues.
- Classification based on modified MICA shows high quality results in lesion/tumor detection. However, it suppresses the majority information before wavelet reconstruction. In some occasions, this will adversely affect the normal brain tissue analysis. MICA can be refined to give equal priority for normal and abnormal tissues in multispectral analysis.

- Expansion of multispectral data with other specific MRI sequences such as diffusion weighted images/perfusion images or sequences with different parameter settings in diagnostic studies can extend the applications to analysis of several brain diseases.
- The newly proposed methods in this thesis can be extended to other application domains of multispectral imaging such as remote sensing, fluorescence microscopy, biometric pattern recognition and so on.
- Real environment artifacts like gaussian noise, illumination effects etc., were not considered in the analysis in this thesis, which can be explored in future as a separate research area.

REFERENCES

1. T. Taxt and A. Lundervold, Multispectral analysis of the brain using magnetic resonance imaging, *IEEE Trans. Med. Imaging*, 13(3):470-481, 1994.

2. J. T. Bushberg, J. A Seibert, E. M. Leidholdt, and J. M. Boone, *The Essential Physics of Medical Imaging,* 2nd Edition, Lippincott Williams & Wilkins, 2001.

3. P. A. Rinck, *Magnetic Resonance in Medicine*, Blackwell Publishers, 2001.

4. Y. Kvinnsland, N. Brekke, T.M. Taxt, and R. Gruner, Multispectral analysis of multimodal images, *Acta Oncol.,* 48(2): 277-284, 2009.

5. Y.C. Ouyang, H.M. Chen, J.W. Chai, C.C. Chen, Clayton C.C. Chen, S.K Poon, C.W. Yang, and S.K. Lee, Independent component analysis for magnetic resonance image analysis, *EURASIP J. Adv. Signal Process*, 2008:780656, 2008.

6. R. He, S. Datta, B.R. Sajja, and P.A. Narayana, Generalized fuzzy Clustering for segmentation of multi-spectral magnetic resonance images, *Comput. Med. Imaging Graph,* 32(5): 353-366, 2008.

7. L.P. Clarke, R.P. Velthuizen, M.A. Camacho, J.J. Heine, M. Vaidyanathan, L.O. Hall, R.W. Thatcher, and M.L. Silbiger, Mri segmentation: methods and applications, *Magn. Reson. Imaging,* 13(3):343-368, 1995.

8. C .Valdés Hernández Mdel, P.J. Gallacher, M.E. Bastin, N.A. Royle, S.M. Maniega, I.J. Deary, and J.M. Wardlaw, Automatic segmentation of brain white matter and white matter lesions in normal aging: comparison of five multispectral techniques, *Magn. Reson. Imaging*, 30(2): 222-229, 2012.

9. J. P. Hornak, The Basics of MRI, [online book], 2004. Website: http://www.cis.rit.edu/htbooks/mri/

10. Website: https://wiki.engr.illinois.edu/display/BIOE414/MRI+Introduction

11. Website:http://www.nhs.uk/Conditions/MRI-scan/Pages/How-does-it-work.aspx

12. D. B. Twieg, The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods, *Med. Phys*., 10:610–621, 1983.

13. R. Bitar, G. Leung, R. Perng, S. Tadros, A. R. Moody, J. Sarrazin, C.McGregor, M . Christakis, S. Symons, A. Nelson, and T. P. Roberts, MR pulse sequences: what every radiologist wants to know but is afraid to ask, *Radiographics,* 26(2):513-37, 2006

14. Website: http://www.medicalphysicist.co.uk/chapter3.pdf

15. C. Westbrook, MRI *at a glance*, Oxford, England: Blackwell Science, 2002.

16. M. A. Brown and R. C. Semelka, MR imaging abbreviations, definitions, and descriptions: a review, *Radiology*, 213:647–662, 1999.

17. R. H. Hashemi, W. G. Bradley, D. Y. Chen, J. E. Jordan, J. A. Queralt, A. E. Cheng, and J. N. Henrie, Suspected multiple sclerosis: MR imaging with a thin-section fast flair pulse sequence, *Radiology,*196(2):505-510, 1995.

18. K.R.R. Krishnan and J.R. MacFall **,** CAMRD Project 268**,** *Duke Silvio Conte Center.*
http://camrd4.mc.duke.edu/camrdgallery/CAMRDResearch/camrd268.html

19. D. L. Pham, C. Xu, and J. L. Prince, Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2:315-337, 2000.

20. C.A.Cocosco, V. Kollokian, R.K.-S. Kwan, and A.C. Evans, Brainweb: online interface to a 3D MRI simulated brain database, *Neuroimage,* 5(4) S425, 1997. http://www.bic.mni.mcgill.ca/brainweb

21. H. K. Ramapriyan, *Survey of Image Analysis Software*, Eastern Regional Remote Sensing Applications Center, 1980.

22. M. W. Vannier, R. L. Butterfield, D. Jordan, W. A. Murphy, R. G. Levitt, and M. Gado, Multispectral analysis of magnetic resonance images, *Radiology*, 154(1):221-224, 1985.

23. J. K. Gohagan, E. L. Spitznagel, W. A. Murphy, M. W. Vannier, W. T. Dixon, D. J. Gersell, S. L. Rossnick, W. G. Totty, J. M. Destouet, D. L. Rickman, T. A. Spraggins, and R. L. Butterfield, Multispectral analysis of MR images of the breast, *Radiology*, 163(3):703-7, 1987.

24. X. Llado, A. Oliver, M. Cabezas, J. Freixenet, J.C. Vilanova, A. Quiles, L. Valls, L. Ramio-Torrenta, and A. Rovira, Segmentation of multiple sclerosis lesions in brain MRI: a review of automated approaches, *Information Sciences*, 186(1):164-185, 2012.

25. R. de Boer, H.A. Vrooman, M.A. Ikram, M.W. Vernooij, M.M. Breteler, A. van der Lugt, and W.J. Niessen, Accuracy and reproducibility study of automatic MRI brain tissue segmentation methods, *Neuroimage*, 51(3):1047-56, 2010.

26. C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

27. D.F. Specht, Probabilistic Neural Networks, *Neural Networks,* 3(1):109-118, 1990.

28. V.N. Vapnik, *Statistical learning theory*, John Wiley and Sons Inc, New York, 1998.

29. J.W. Chai, C. Chi-Chang Chen, C.M. Chiang, Y.J. Ho, H.M. Chen, Y.C. Ouyang, C.W. Yang, S.K. Lee, and C.I. Chang, Quantitative analysis in clinical applications of brain MRI using independent component analysis coupled with support vector machine, *J. Magn. Reson. Imaging*, 32(1):24-34, 2010.

30. N. Abdullah, U.K. Ngah, and S.A. Aziz, Image classification of brain MRI using support vector machine, In *Proc. IEEE Conf. on Imaging Systems and Techniques (IST), Malaysia*, pp. 242 – 247, 2011.

31. E. Eyal, B. N. Bloch, N. M. Rofsky, E. Furman-Haran, E. M. Genega, R. E. Lenkinski, and H. Degani, Principal component analysis for dynamic contrast-enhanced MRI in human prostate cancer, *Invest. Radiol*. 45(4):174–181, 2010.

32. S. Alzubi, N. Islam, and M. Abbod, Multiresolution analysis using wavelet, ridgelet and curvelet transforms for medical imaging segmentation, *Int. J. of Biomed. Img.*, 2011, Article ID 136034, 18 pages, 2011.

33. T.Nakai, S.Muraki, E.Bararinao, Y.Miki, Y.Takehara, K.Matsuo, C.Kato, H.Sakabara, and H.Isoda, Application of independent component analysis to magnetic resonance imaging for enhancing the contrast of gray matter and white matter, *Neuroimage*, 21(1):251–260, 2004.

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

35. K.J. Friston, J. Ashburner, C. Frith, J.B. Poline, J. D. Heather, and R.S.J. Frackowiak, Spatial registration and normalization of images, *Hum. Brain. Mapp.,* 2:165-189, 1995.

36. J. Ashburner and K.J. Friston, Unified segmentation, *Neuroimage*, 26(3):839-51, 2005.

37. SPM Website : http://www.fil.ion.ucl.ac.uk/spm/ , 2000.

38. FSL Website :http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/ ,2000.

39. S.M. Smith, M. Jenkinson, M.W. Woolrich, C.F. Beckmann, T.E.J. Behrens, H. Johansen-Berg, P.R. Bannister, M. De Luca, I. Drobnjak, D.E. Flitney, R. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J.M. Brady, and P.M. Matthews, Advances in functional and structural MR image analysis and implementation as FSL, *Neuroimage*, 23 Suppl 1:S208-19, 2004.

40. S.M. Smith, Fast robust automated brain extraction, *Hum. Brain. Mapp.*, 17(3):143-55, 2002.

41. A. M. Dale., B. Fischl, and M.I. Sereno, Cortical surface-based analysis: segmentation and surface reconstruction, *Neuroimage,* 9(2):179-94, 1999.

42. X. Han and B. Fischl, Atlas renormalization for improved brain MR image segmentation across scanner platforms, *IEEE Trans. Med. Imaging,* 26(4):479-486, 2007.

43. F. Sègonne, A.M. Dale, E. Busa, M.Glessner, D. Salat, H. K. Hahn and B. Fischl, A hybrid approach to the skull stripping problem in MR, *Neuroimage,* 22(3):1060-1075, 2004.

44. Y. Zhang, M. Brady, and S. Smith, Segmentation of brain MR images through a hidden markov random field model and the expectation- maximization algorithm, *IEEE Trans. Med. Imaging*, 20(1):45-57, 2001.

45. O. Tsang, A. Gholipour, N. Kehtarnavaz, K. Gopinath, R. Briggs, and I. Panahi, Comparison of tissue segmentation algorithms in neuroimage

analysis software tools, In *Conf. Proc. IEEE Eng. Med. Biol. Soc., 2008:3924-8*, *2008.*

46. D. W. Shattuck, G. Prasad, M. Mirza, K. L. Narr, and A. W. Toga, Online resource for validation of brain segmentation methods, *Neuroimage,* 45(2):431-9, 2009.

47. C. H. Mortiz, V. M. Haughton, D. Cordes, M. Quigley, and M. E. Meyerand, Whole brain functional MR imaging activation from finger tapping task examined with independent component analysis, *AJNR Am. J. Neuroradiol.,* 21(9):1629-35, 2000.

48. R. N. Bracewell, *The Fourier Transform and Its Applications*, 3rd Edition, McGraw-Hill, New York, 1999.

49. S. Chaplot, L. M. Patnaik, and N. R. Jagannathan, Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network, *Biomedical Signal Processing and Control*, 1(1): 86-92, 2006.

50. S. Mallat, *A Wavelet Tour of Signal Processing*, 3rd Edition, The Sparse Way Academic Press, 2008.

51. E.S. A. Dahshan, T. Hosny, and A.B. M. Salem, Hybrid intelligent techniques for MRI brain images classification, *Digital Signal Processing*, 20(2): 433-441, 2010.

52. J. Ma and G. Plonka, The curvelet transform: a review of recent applications, *IEEE Sig. Process. Mag.*, 27(2):118-133, 2010.

53. H. Han and X. Li, Multi-resolution independent component analysis for high-performance tumor classification and biomarker discovery, *BMC Bioinformatics*, 12 (Suppl. 1:S7), 2011.

54. K. Müller, G. Lohmann, S. Zysset, and D. Y. von Cramon, Wavelet statistics of functional MRI data and the general linear model, J. *Magn. Reson. Imaging*, 17(1):20-30, 2003.

55. S. Kaewpijit, J. L. Moigne, and T. EL-Ghazawi, Automatic reduction of hyperspectral imagery using wavelet spectral analysis, *IEEE Trans. Geosci. Remote Sensing*, 41(4): 863-871, 2003.

56. A. Pitiot, H. Delingette, P. M. Thompson, and N. Ayache, Expert knowledge-guided segmentation system for brain MRI, *Neuroimage,* 23 Suppl 1:S85-96, 2004

57. P. Golland, W. E. Grimson, M. E. Shenton, and R. Kikinis, Detection and analysis of statistical differences in anatomical shape, *Med. Image Anal.,* 9(1):69-86, 2005.

58. S. Powell, V.A. Magnotta, H. Johnson, V. K. Jammalamadaka, R.Pierson, and N. C. Andreasen, Registration and machine learning-based automated segmentation of subcortical and cerebellar brain structures, *Neuroimage,* 39(1):238-47, 2008.

59. W. Messen, R. Wehrens, and L. Buydens, Supervised kohonen networks for classification problems, *Chemometrics and Intelligent Laboratory Systems*, 83:99-113, 2006.

60. K. Yamashita, T. Yoshiura, H. Arimura, F. Mihara, T. Noguchi, A. Hiwatashi, O. Togao, Y. Yamashita, T. Shono, S. Kumazawa, Y. Higashida, and H. Honda, Performance evaluation of radiologists with artificial neural network for differential diagnosis of intra-axial cerebral tumours on MR images, *AJNR Am J Neuroradiol.,* 29(6):1153-8, 2008.

61. E. S. A. El-Dahshan, A. B. M. Salem, and T. H. Younis, A hybrid technique for automatic MRI brain images classification, *Studia Univ, Babes Bolyai. Informatica*, *54*(1), 55-67, 2009.

62. T. R. Jensen and K. M. Schmainda**,** Computer-aided detection of brain tumour invasion using multiparametric MRI, *J. Magn. Reson. Imaging,* 30(3):481–9, 2009.

63. AmirEhsan Lashkari, A neural network based method for brain abnormality detection in MR images using Gabor wavelets, *Intl. J. Comput. Appl.*, 4(7):9-15, 2010.

64. D. M. Joshi, N. K. Rana, and V. M. Misra, Classification of brain cancer using artificial neural network, *Intl. Conf. Electronic Comput. Technol,* pp: 112-116, IEEE 2010.

65. Y. Zhang, Z. Dong, L. Wu, and S. Wang, A hybrid method for MRI brain image classification, *Expert Syst. Appl.*, 38(8):10049-10053, 2011.

66. G. Z. Li, J. Yang, C. Z. Ye, and D. Y. Geng, Degree prediction of malignancy in brain glioma using support vector machines, *Comput. Biol. Med.*, 36(3):313-25, 2006.

67. J. Luts, A. Heerschap, J. A. Suykens, and S. Van Huffel, A combined MRI and MRSI based multiclass system for brain tumour recognition using LS-SVMs with class probabilities and feature selection, *Artif. Intell. Med.,* 40(2):87-102, 2007.

68. S. Selvaraj, S. Thamaraiselvi, D. Selvathi, and L. Gewali, Brain MRI slices classification using least squares support vector machine, *International Journal of Intelligent Computing and Medical Sciences in Image Processing(ICMED)*, 1:21-33, 2007.

69. S. Chandra, R. Bhat, H. Singh, and S. Chauhan, Detection of brain tumors from MRI using Gaussian RBF kernel based support vector machine*, International Journal of Digital Content Technology and its Applications,* 1:46-51, 2009.

70. A. Hamilton-Wright, W. Daniel, and R. Hamid, Fuzzy classification using pattern discovery, *IEEE Trans. on Fuzzy Systems,* 15(5):772-83, 2007.

71. E. I. Zacharaki, S. Wang, S. Chawla, D. Soo Yoo, R. Wolf, E. R. Melhem, and C. Davatzikos, Classification of brain tumour type and grade using MRI texture and shape in a machine learning scheme, *Magn. Reson. Med.*, 62 (6)1609-1618, 2009.

72. S. Ruan, N. Zhang, Q. Liao, and Y. Zhu, Image fusion for following-up brain tumour evolution*, Proc. IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, Chicago, pp. 281–4, 2011.

73. Y. Zhang and L. Wu, An MR brain iImages classifier via principal component analysis and kernel support vector machine, *Progress in Electromagnetics Research*, 130:369-388, 2012.

74. Mubashir Ahmad, Mahmood ul-Hassan, Imran Shafi, and Abdelrahman Osman, Classification of tumors in human brain MRI using wavelet and

support vector machine, *IOSR Journal of Computer Engineering (IOSRJCE)*, 8(2):25-31, 2012.

75. P. Georgiadis, D. Cavouras, J. Kalatzis, A. Daskalakis, G. C. Kagadis, K. Sifaki, M. Malamas, G. Nikiforidis, and E. Solomou, Improving brain tumor characterization on MRI by probabilistic neural networks and non-linear transformation of textural features, *Comput. Methods Programs Biomed.,* 89(1):24-32, 2008.

76. M. Ibrahiem and S. Ramakrishnan, On the application of various probabilistic neural networks in solving different pattern classification problems, World *Applied Sciences Journal,* 4(6):772-780, 2008.

77. J. Han and M. Kamber, Data *Mining: Concepts and Techniques*, Elsevier, 2009.

78. L. Breiman, Bagging predictors, *Machine Learning*, 24(2):123-140,1996.

79. H. C. Kim, S. Pang, H. M. Je, D. Kim, and S. Y. Bang, Support vector machine ensemble with bagging*, Lecture Notes in Computer Science*, 2388:397-408, 2002.

80. L. R. Schad, S. Bluml, and I. Zuna, MR tissue characterization of intracranial tumors by means of texture analysis*, Magn. Reson. Imaging*, 11(6):889-896, 1993.

81. W. E. Phillips, R. P. Velthuizen, S. Phuphanich, L. O. Hall, L. P. Clarke, and M. L. Silbiger, Application of fuzzy c-means segmentation technique for tissue differentiation in MR images of a hemorrhagic glioblastoma multiforme, *Magn. Reson. Imaging,* 13(2):277-90, 1995.

82. M. Vaidyanathan, L.P. Clarke, R.P. Velthuizen, S. Phuphanich, A. M. Bensaid, L.O. Hall, J. C. Bezdek, H. Greenberg, A. Trotti, and M. Silbiger**,** Comparison of supervised MRI segmentation methods for tumor volume determination during therapy, *Magn. Reson. Imaging,* 13(5):719–28, 1995.

83. M. C. Clark, L. O. Hall, D. B. Goldgof, R. Velthuizen, F. R. Murtagh, and M. S. Silbiger, Automatic tumor segmentation using knowledge-based techniques, *IEEE Trans. Med. Imaging,* 17(2):187–201, 1998.

84. Dzung L. Pham and Jerry L. Prince, An adaptive fuzzy c-means algorithm for image segmentation in the presence of intensity inhomogeneties, *Pattern Recognition Letters,* 20(1):57-68, 1999.

85. A. Liew and H. Yan, An adaptive spatial clustering algorithm for 3D MR image segmentation, *IEEE Trans. Med. Imag.*, 22(9):1063-75, 2003.

86. L. M. Fletcher-Heath, L. O. Hall, D. B. Goldgof, and F. R. Murtagh, Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. *Artif. Intell. Med*, 21(1):43-63, 2001.

87. N. Moon, E. Bullitt, K. Van Leemput, and G. Gerig, Model based brain and tumor segmentation, *Int. Conf. on Pattern Recognition,* 1:528-531, 2002.

88. F. Zhu and J. Tian, Modified fast marching and level set method for medical image segmentation, *J. Xray Sci. Technol.,* 11(4):193-204, 2003.

89. Y. Li and Z. Chi, MR brain image segmentation based on self-organizing map network, *International Journal of Information Technology*, 11(8):45-53, 2005.

90. S. Deorah, C. F. Lynch, Z. A. Sibenaller, and T. C. Ryken, Trends in brain cancer incidence and survival in the United States: surveillance, epidemiology, and end results program, 1973 to 2001, *Neurosurg Focus*, 20(4):E1, 2006.

91. H. Zaidi, T. Ruest, F. Schoenahl, and M.L. Montandon, Comparative assessment of statistical brain MR image segmentation algorithms and their impact on partial volume correction in PET*, Neuroimage*, 32(4):1591-607, 2006.

92. M. Martin-Landrove and R. Villalta, Brain tumor image segmentation using neural networks, *Proc. of International Society of Magnetic Resonance in Medicine*, 14:1610, 2006.

93. H. Khotanlou, O. Colliot, and I. Bloch, Automatic brain tumor segmentation using symmetry analysis and deformable models, Int*. Conf. On Advances in Pattern Recognition,* pp.198-202, 2007.

94. R. Unnikrishnan, C. Pantofaru, and M. Hebert, Toward objective evaluation of image segmentation algorithms, *IEEE Trans. Pattern. Anal. Machine Intell.*, 29(6):929-44, 2007.

95. N. Ray, R. Greiner, and A. Murtha, Using Symmetry to Detect Abnormalities in Brain MRI, *Computer Society of India Communications*, 31(19), pp 7--10, January 2008.

96. C. H. Lee, S. Wang, A. Murtha, M. R. G. Brown, and R. Greiner, Segmenting brain tumors using pseudo conditional random fields, *Proc. of the 11th Int. Conf. on Medical Image Computing and Computer Assisted Intervention (MICCAI),* 5241:359-66, 2008.

97. R. Khayati, M. Vafadust, F. Towhidkhah, and M. Nabavi, Fully automatic segmentation of multiple sclerosis lesions in brain MR FLAIR images using adaptive mixtures method and markov random field model, *Comput. Biol. Med.,* 38(3): 379-390, 2008.

98. J. J. Corso, E. Sharon, S. Dube, S. El-Saden, U. Sinha, and A. Yuille, Efficient multilevel brain tumor segmentation with integrated Bayesian model classification, *IEEE Trans. Med. Imag*., 27(5): 629 - 640, 2008.

99. P. Anbeek, K. L. Vincken, M. A. Viergever, Automated MSLesion segmentation by K-Nearest Neighbour classification, *The MIDAS online Journal*, 2008. http://hdl.handle.net/10380/1448

100. Z. Ma, J. M. Tavares, and R. M. Natal Jorge, A review on the current segmentation algorithms for medical images, *Int. Conf. on Imaging Theory and Applications*, 135-40, 2009.

101. R. de Boer, H. A. Vrooman, M. A. Ikram, M. W. Vernooij, M. M. Breteler, A. van der Lugt, and W. J. Niessen, Accuracy and reproducibility study of automatic MRI brain tissue segmentation methods, *Neuroimage*, 51(3):1047-56, 2010.

102. C. W. Kanaly, D. Ding, A. I. Mehta, A. F. Waller, I. Crocker, A. Desjardins, D. A. Reardon, A. H. Friedman, D. D. Bigner, and J. H. Sampson, A novel method for volumetric MRI response assessment of enhancing brain tumors, *PloS One,* 6(1): e16031, 2011.

103. T. W. Cheng, D. B. Goldgof, and L. O. Hall, Fast fuzzy clustering, *Fuzzy Sets and Systems*, 93(1):49-56, 1998.

104. A. O. Boudraa, S. M. Dehak, Y. M. Zhu, C. Pachai, Y. G. Bao, and J. Grimaud, Automated segmentation of multiple sclerosis lesions in multispectral MR imaging using fuzzy clustering, *Comput. Biol. Med.*, 30(1): 23–40, 2000.

105. R. N. Dave and S. Sen, Robust fuzzy clustering of relational data, *IEEE Trans. Fuzzy Systems,* 10(6):713-27, 2002.

106. J. F. Kolen and T. Hutcheson, Reducing the time complexity of the fuzzy C-means algorithm, *IEEE Trans. on Fuzzy Systems*, 10(2):263- 267, 2002.

107. M. M. Khalighi, H. Soltanian-Zadeh, and C. Lucas, Unsupervised MRI segmentation with spatial connectivity, *Proc. SPIE* 4684, Medical Imaging 2002: Image Processing, 1742, 2002.

108. S. Eschrich, J. Ke, L. O. Hall, D. B. Goldgof, Fast accurate fuzzy clustering through data reduction, *IEEE Trans. on Fuzzy Systems*, 11(2):262-70, 2003.

109. D. Q. Zhang and S. C. Chen, A novel kernelized fuzzy c-means algorithm with application in medical image segmentation**,** *Artif Intell Med.* 32(1):37-50, 2004.

110. S. R. Kannan, Segmentation of MRI using new unsupervised fuzzy c-means algorithm, *Graphics, Vision and Image Processing Journal,* 5(2):17-23, 2005.

111. S. Murugavalli and V. Rajamani, A high speed parallel fuzzy c-mean algorithm for brain tumor segmentation, *Bioinformatics and Medical Engineering Journal,* 6(1):29-34, 2006.

112. W. Dou, S. Ruan, Y. Chen, D. Bloyet, and J. Constans, A framework of fuzzy information fusion for the segmentation of brain tumor tissues on MR images, Image *and Vision Computing,* 25(2):164-71, 2007.

113. C. F. Juang, S. H. Chiu, S. W. Chang, A self organizing TS-type fuzzy network with support vector learning and its application to classification problems, *IEEE Trans. on Fuzzy Systems*, 15(5):998-1008, 2007.

114. S. R. Kannan, A new segmentation system for brain MR images based on fuzzy techniques, *Applied Soft Computing*, 8(4):1599-606, 2008.

115. X. Zhou, Q. Shen, and L. Liu, New two-dimensional fuzzy c-means clustering algorithm for image segmentation, *J. Central South Univ. Technol.*, 15(6):882-887, 2008.

116. S. Ramakrishnan, M. M. Ibrahiem, and E. l. Emary, Classification brain MR images through a fuzzy multiwavelets based GMM and probabilistic neural networks, *Telecommun. Syst., Springer Sci,* 46(3):245-252, 2010.

117. X. Yang and B. Fei, A multiscale and multiblock fuzzy c-means classification method for brain MR images, *Med. Phys.,* 38(6):2879–2891, 2011.

118. H. Kang, A. Pinti, A. Taleb-Ahmed, and X. Zeng, An intelligent generalized system for tissue classification on MR images by integrating qualitative medical knowledge, *Biomed. Signal Processing and Control,* 6(1):21–26, 2011.

119. Iraky khalifa, Aliaa Youssif, and Howida Youssry, MRI Brain Image Segmentation based on Wavelet and FCM Algorithm, *International Journal of Computer Applications*, 47(16):32-39, 2012.

120. J. C. Bezdec, *Pattern Recognition with Fuzzy Objective Function Algorithms,* Plenum Press, New York, 1981.

121. G. Gerig, J. Martin, R. Kikinis, O. Kubler, M. Shenton, and F. A. Jolesz, Unsupervised tissue type segmentation of 3D dual echo MR head data. *Image and Vision Computing*, 10(6):349-360, 1992

122. T. Kapur, W. E. L. Grimson, W. M. Wells, and R. Kikinis, Segmentation of brain tissue from magnetic resonance images. *Medical Image Analysis*, 1(2):109-127, 1996.

123. Y. F. Tsai, I. J. Chiang, Y. C. Lee, C. C. Liao, and K. L. Wang, Automatic mri meningioma segmentation using estimation maximization, *Conf Proc IEEE Eng Med Biol Soc.,* 3:3074-7*,* 2005.

124. B. S. Everitt, Cluster analysis: a brief discussion of some of the problems. *The British Journal of Psychiatry*, 120(555):143-145, 1972.

125. J. R. Mitchell, S. J. Karlik, D. H. Lee, and A. Fenster, Computer-assisted identification and quantification of multiple sclerosis lesions in mr imaging

volumes in the brain, *Journal of Magnetic Resonance Imaging,* 4(2):197-208, 1994.

126. M. Ozkan, B. M. Dawant, and R. J. Maciunas., Neural network-based segmentation of multi-modal medical images: a comparative and prospective study, *IEEE Trans. Med. Imaging,* 12(3):534-544,1993.

127. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd Edition, Wiley, 2001.

128. S. Ruan, S. Lebonvallet, A. Merabet, and J. M. Constans, Tumor segmentation from a multispectral MRI images by using support vector machine classification, *4th IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, 2:1236–9, 2007

129. R. Verma, E. Zacharaki, Y. Ou, H. Cai, S. Chawla, S. K. Lee, E. R. Melhem, R. Wolf, and C. Davatzikos, Multiparametric tissue characterization of brain neoplasms and their recurrence using pattern classification of MR images, *Acad. Radiol,* 15(8):966–77, 2008.

130. D. E. Gustafson and W. C. Kessel, Fuzzy clustering with a fuzzy covariance matrix, *Proc. IEEE Conf. Decision Contr.,* 761-766, 1978.

131. D. Kroon, E. van Oort, and K. Slump, Multiple sclerosis detection in multispectral magnetic resonance images with principal components analysis, *3D Segmentation in the Clinic: A Grand Challenge II: MS lesion segmentation,* pp. 1-14, 2008.

132. Z. Lao, D. Shen, D. Liu, A. F. Jawad, E. R. Melhem, L. J. Launer, R. N. Bryan, and C. Davatzikos, Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine, *Acad. Radiol.,* 15(3):300-13, 2008.

133. J. Lecoeur, J.C. Ferré, and C. Barillot, Optimized supervised segmentation of MS lesions from multispectral MRIs, *Work. Med. Image Anal. Mult. Scler*, pp. 5–14, 2009.

134. S. Ozer, D. L. Langer, X. Liu, M. A. Haider, T. H. van der Kwast, A. Evans, Y. Yang, M. N. Wernick, and I. S. Yetik, Supervised and unsupervised methods

for prostate cancer segmentation with multispectral MRI, *Med. Phys.*, 37(4):1873-83, 2010.

135. N. Zhang, S. Ruan, S. Lebonvallet, Q. Liao, and Y. Zhu, Kernel feature selection to fuse multi-spectral MRI images for brain tumor segmentation, *Computer Vision and Image Understanding,* 115(2):256-269, 2011.

136. C. Hernández Mdel, K. J. Ferguson, F. M. Chappell, and J. M. Wardlaw, New multispectral MRI data fusion technique for white matter lesion segmentation: method and comparison with thresholding in FLAIR images, *Eur. Radiol*., 20(7):1684–91, 2010.

137. Analyze package, Mayo Clinic, KS, USA. website: http://www.analyzedirect.com/.

138. J. M. Bland and D. G. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, *Lancet,* 1: 307–310, 1986.

139. D. García-Lorenzo, S. Francis, S. Narayanan, D. L. Arnold, and D. L. Collins, Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging, *Med. Image Anal.*, 17(1):1–18, 2013.

140. J. F. Cardoso, Eigenstructure of the 4th-order cumulant tensor with application to the blind source separation problem, *Proc. ICASSP '89*, pp.2109–2112, 1989.

141. C. Jutten and J. Herault, Blind separation of sources part I: An adaptive algorithm based on neuromimetic architecture, *Signal Processing*, 24(1):1–10, 1991.

142. P. Common, Independent component analysis, a new concept?, *Signal Processing, 36(3):287–314, 1994.*

143. A. J. Bell and T. J. Sejnowski*,* An information-maximization approach to blind separation and blind deconvolution*, Neural Comput., 7(6):1129–1159, 1995.*

144. A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons Inc, New York, 2001.

145. C. J. James and C. W. Hesse, Independent component analysis for biomedical signals, *Physiol. Meas.*, 26(1):R15-39, 2005,

146. Y. W. Chen and D. Sugiki, Segmentation of MR images using independent component analysis, *Lecture Notes in Artificial Intelligence*, Springer, LNAI 4252: 63-69, 2006.

147. T. Tateyama, Z. Nakao, and Y. Chen, Classification of Brain Matters in MRI by Kernel Independent Component Analysis, *IIH-MSP IEEE Computer Society*, pp. 713-716, 2008.

148. F. R. Bach and M. I. Jordan, Kernel independent component analysis, Journal *of Machine Learning Research*, 3:1–48, 2002.

149. Wenlu Yang, Xinyun Chen, Hong Xie, and Xudong Huang, ICA-Based Automatic Classification of Magnetic Resonance Images from ADNI Data, *LSMS/ICSEE,* 6330:340-347, 2010.

150. Y.C. Ouyang, H.M. Chen, J.W. Chai, C.C. Chen, S.K. Poon, C.W. Yang, S.K. Lee, and C.I. Chang, Band expansion-based over-complete independent component analysis for multispectral processing of magnetic resonance images, *IEEE Trans. Biomed. Eng.*, 55(6):1666-77, 2008.

151. J. W. Chai, S. K. Lee, Clayton C. C. Chen, H. M. Chen, and Y.C. Ouyang, Independent component analysis in magnetic resonance imaging, *SPIE Newsroom*, 2009. https://spie.org/x34797.xml?ArticleID=x34797

152. C. Bauer, F. J. Theis, W. Baumler, EW. Lang, Local features in biomedical image clusters extracted with independent component analysis, *Proc. Int. Joint Conf. Neural Networks (IJCNN)*, 1: 81-84, 2003.

153. J. Herault and C. Jutten, Space or time processing by neural network models, *Proc. AIP Conference: Neural Networks for Computing*, 151, (New York: AIP), 1986.

154. A. Hyvarinen and E. Oja, A Fast fixed-point algorithm for independent component analysis, *Neural Computation,* 9(7):1483-1492, 1997

155. D. T. Pham, P. Garrat, and C. Jutten, Separation of a mixture of independent sources through a maximum likelihood approach, *Proc. EUSIPCO*, pp. 771–774, 1992.

156. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 3rd Edition, 1991.

157. J. F. Cardoso and A. Souloumiac, Jacobi angles for simultaneous diagonalization, *SIAM J. Matrix Anal. Appl.* 17(1):161-164, 1996.

158. A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, *IEEE Trans. on Neural Networks*, 10(3):626–634, 1999.

159. F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz, The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data, *Remote Sens. Environ.*, 44 (2–3):145-163, 1993.

160. T. Achalakul, S. Taylor, A distributed spectral-screening PCT Algorithm, *J. Parallel Distrib. Comput.* 63 (3):373-384, 2003.

161. H. Han, A high performance profile-biomarker diagnosis for mass spectral profiles, *BMC Syst. Biol.,* 5 (2S5), 2011.

162. V. W. Medha, M. P. Pradeep, and K. A. Hemant, Image registration techniques: an overview, *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2(3), 2009.

163. I. Daubechies, *Ten Lectures on Wavelets,* SIAM, USA, 1992.

164. P. Paclík and R. P. W. Duin, Dissimilarity-based classification of spectra: computational issues. *Real Time Imaging,* 9(4):237–244, 2003.

165. C. Ambroise and G. J. McLachlan, Selection bias in gene extraction on the basis of microarray gene-expression data, *Proc. Natl. Acad. Sci. USA,* 99(10):6562-6566, 2002.

166. A. Ebel, A. A. Maudsley, and N. Schuff, Correction of local B0 shifts in 3D EPSI of the human brain at 4 T, *Magn Reson Imaging*, 25(3):377-380, 2007.

167. B. Alfano, A. Brunetti, E. M. Covelli, M. Quarantelli, M. R. Panico, A. Ciarmiello, and M. Salvatore, Unsupervised, automated segmentation of the normal brain using a multispectral relaxometric magnetic resonance approach, *Magn Reson Med*, 37(1): 84–93, 1997.

168. R. P. Velthuizen, L. P. Clarke, S. Phuphanich, L. O. Hall, A. M. Bensaid, J. A. Arrington, H. M .Greenberg, and M. L. Silbiger, Unsupervised measurement of brain tumor volume on MR images, *J. Magn. Reson. Imaging*, 5(5):594–605, 1995.