7-26-2011

An Information Theoretic Approach For Feature Selection And Segmentation In Posterior Fossa Tumors

Shaheen Ahmed Ms.

To the University Council:

The Dissertation Committee for Shaheen Ahmed certifies that this is the final approved version of the following electronic dissertation: “An Information Theoretic Approach For Feature Selection And Posterior Fossa Tumor Segmentation.”

_____
Khan M. Iftekharuddin, Ph.D.
Major Professor

We have read this dissertation and recommend its acceptance:

_____
Ebenezer O. George, Ph.D.

_____
David J. Russomanno, PhD.

_____
Robert J. Ogg, Ph.D.

Accepted for the Graduate Council:

_____
Karen D. Weddle-West, Ph.D.
Vice Provost for Graduate Programs

AN INFORMATION THEORETIC APPROACH FOR FEATURE SELECTION AND SEGMENTATION IN POSTERIOR FOSSA TUMOR

by

Shaheen Ahmed

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Electrical and Computer Engineering

The University of Memphis

August 2011

To my wonderful parents

ACKNOWLEDGMENT

I am thankful to Dr. Khan for his support for supervising me during my thesis. I appreciate his patience and guidance while discussing research objectives. I would like to thanks Dr. George for helping me with statistics modeling and the knowledge which I gained during his statistical classes. I express my gratitude to Dr. Vossough for providing us with the datasets and providing his valuable time for analyzing the MR images. I also am thankful to Dr. Ogg and Dr. Russomanno for their valuable suggestions and comments.

I give my special thanks to my parents for encouraging and guiding me at every step of my life. Among my friends, I would like to thank Rahul Sharma who was also there to help me when needed. Also would like to thank Keith, Paul and other members of ISIP lab. Finally, I would like to thank Ms. Becky Ward for being there for me whenever I needed her.

ABSTRACT

Ahmed, Shaheen. Ph.D. The University of Memphis. August 2011. Information Theoretic Approach for Feature Selection and Segmentation in Posterior Fossa Tumor. Major Professor: Khan M. Iftekharuddin, Ph.D.

Posterior fossa (PF) is a type of brain tumor located in or near brain stem and cerebellum. About 55%- 70% pediatric brain tumors arise in the posterior fossa, compared with only 15%- 20% of adult tumors. For segmenting PF tumors we should have features to study the characteristics of tumors. In literature, different type of texture features such as Fractal Dimension (FD) and Multifractional Brownian Motion (mBm) have been exploited for measuring randomness associated with brain and tumor tissue structures, and the varying appearance of tissues in magnetic resonance images (MRI). For selecting best features techniques such as neural network and boosting methods have been exploited. However, neural network cannot describe about the properties of texture features. We explore methods such as information theoretic methods which can perform feature selection based on properties of texture features.

The primary contribution of this dissertation is investigating efficacy of different image features such as intensity, fractal texture, and level-set shape in segmentation of posterior fossa (PF) tumor for pediatric patients. We explore effectiveness of using four different feature selection and three different segmentation techniques respectively to discriminate tumor regions from normal tissue in multimodal brain MRI. Our result suggest that Kullback-Leibler divergence (KLD) measure for feature ranking and selection and Expectation Maximization (EM) algorithm for feature fusion and tumor segmentation offer the best performance for the patient data in this study.

To improve segmentation accuracy, we need to consider abnormalities such as cyst, edema and necrosis which surround tumors. In this work, we exploit features which describe properties of cyst and technique which can be used to segment it. To achieve this goal, we extend the two class KLD techniques to multiclass feature selection techniques, so that we can effectively select features for tumor, cyst and non tumor tissues. We compute segmentation accuracy by computing number of pixels segmented to total number of pixels for the best feature.

For automated process we integrate inhomogeneity correction, feature selection using KLD and segmentation is an integrated EM framework. To validate results we have used similarity coefficients for computing the robustness of segmented tumor and cyst.

PREFACE

This dissertation follows the guideline of IEEE Transaction on Information Technology in Biomedicine 2011. This dissertation was also submitted to IEEE Transaction on Information Technology in Biomedicine 2011.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

1.1. Dissertation Overview

The adult body normally forms new cells only when they are needed to replace old or damaged ones. Infants and children form new cells to complete their development in addition to those needed for repair. A tumor develops if normal or abnormal cells multiply when they are not needed. A brain tumor is a mass of unnecessary cells growing in the brain. There are two basics kinds of brain tumors such as primary tumors and metastatic tumors. Primary brain tumor starts and tends to stay in the brain. Metastatic brain tumor begins as a cancer elsewhere in the body and spreads in the brain.

Tumors are diagnosed and then named based on a classification system. Most medical centers now use the World Health Organization (WHO) classification system for this purpose. Table 1 shows WHO classification of Central Nervous Tumors.

Table 1.1 WHO Classification of Central Nervous System Tumors [28]

Neuroepithelial tumors

1. Astrocytic tumors [glial tumors--categories I-V, below--may also be subclassified as invasive or non-invasive, although this is not formally part of the WHO system, the non-invasive tumor types are indicated below. Categories in italics are also not recognized by the new WHO classification system, but are in common use.]
    1. Astrocytoma (WHO grade II)
        i. variants: protoplasmic, gemistocytic, fibrillary, mixed
    2. Anaplastic (malignant) astrocytoma (WHO grade III)
        i. *hemispheric*
        ii. *diencephalic*
        iii. *optic*
        iv. *brain stem*
        v. *cerebellar*
    3. Glioblastoma multiforme (WHO grade IV)
        i. variants: giant cell glioblastoma, gliosarcoma
    4. Pilocytic astrocytoma [non-invasive, WHO grade I]

        i. *hemispheric*
        ii. *diencephalic*
        iii. *optic*
        iv. *brain stem*
        v. *cerebellar*
    5. Subependymal giant cell astrocytoma [non-invasive, WHO grade I]
    6. Pleomorphic xanthoastrocytoma [non-invasive, WHO grade I]
2. Oligodendroglial tumors
    1. Oligodendroglioma (WHO grade II)
    2. Anaplastic (malignant) oligodendroglioma (WHO grade III)
3. Ependymal cell tumors
    1. Ependymoma (WHO grade II)
        i. variants: cellular, papillary, epithelial, clear cell, mixed
    2. Anaplastic ependymoma (WHO grade III)
    3. Myxopapillary ependymoma
    4. Subependymoma (WHO grade I)
4. Mixed gliomas
    1. Mixed oligoastrocytoma (WHO grade II)
    2. Anaplastic (malignant) oligoastrocytoma (WHO grade III)
    3. Others (*e.g.* ependymo-astrocytomas)
5. Neuroepithelial tumors of uncertain origin
    1. Polar spongioblastoma (WHO grade IV)
    2. Astroblastoma (WHO grade IV)
    3. Gliomatosis cerebri (WHO grade IV)
6. Tumors of the choroid plexus
    1. Choroid plexus papilloma
    2. Choroid plexus carcinoma (anaplastic choroid plexus papilloma)
7. Neuronal and mixed neuronal-glial tumors
    1. Gangliocytoma
    2. Dysplastic gangliocytoma of cerebellum (Lhermitte-Duclos)
    3. Ganglioglioma
    4. Anaplastic (malignant) ganglioglioma
    5. Desmoplastic infantile ganglioglioma
        i. *desmoplastic infantile astrocytoma*
    6. Central neurocytoma
    7. Dysembryoplastic neuroepithelial tumor
    8. Olfactory neuroblastoma (esthesioneuroblastoma)
        i. variant: olfactory neuroepithelioma
8. Pineal Parenchyma Tumors
    1. Pineocytoma
    2. Pineoblastoma
    3. Mixed pineocytoma/pineoblastoma
9. Tumors with neuroblastic or glioblastic elements (embryonal tumors)
    1. Medulloepithelioma
    2. Primitive neuroectodermal tumors with multipotent differentiation

        i. medulloblastoma
        ii. cerebral primitive neuroectodermal tumor

    3. Neuroblastoma
        i. variant: ganglioneuroblastoma
    4. Retinoblastoma
    5. Ependymoblastoma
10. Tumors of the sellar regions
    1. Pituitary adenoma
    2. Pituitary carcinoma
    3. Craniopharyngioma
    4. Hematopoietic tumors
    5. Primary malignant lymphomas
        i. Plasmocytoma
        ii. Granulocytic sarcoma

11. Germ cell tumors
    i. Germinoma
    ii. Embryonal carcinoma
1. Tumor of meninges
    i. Meningioma
    ii. Atypical meningioma
    iii. Anaplastic meningioma
    iv. Non- meningothelial tumor of meninges
2. Tumor of cranial and spinal neves
    i. Schwannoma
    ii. Neurofibroma
    iii. Malignant peripheral nerve sheath tumor
12. Metastatic tumors
13. Unclassified tumors
14. Cysts and tumor –like lesions

Note in Table 1.1 we can observe that the tumors are graded from 1-4. The grade of a tumor indicates its degree of malignancy. Grade 1 tumors are the least malignant. Grade 2 tumors are the relatively slow growing; they can spread into nearby normal tissue and can recur. Grade 3 tumors are malignant and high grade tumors. Grade 4 is most malignant tumors and can easily grow into surrounding normal tissues.

Nearly 1,500 to 2,000 children in US are affected by brain tumors every year [1]. Such pediatric brain tumor can result from abnormal growth of tissues either in the brain or in other internal organs leading to metastasis of mass in brain. Diseases such as neurofibromatosis, von Hippel-Lindau disease, Li- Fraumeni syndrome and retinoblastoma are all associated with a higher risk of brain tumors in children [2]. Although pediatric brain tumors may originate at any age, children are mostly diagnosed with brain tumors between the ages of three and eight. These tumors can be more or less malignant that may grow rapidly and spread to the spinal cord.

Some examples of childhood brain tumors include Astrocytomas, Glioblastoma multiform (GBM), Ependymomas, Primitive neuroectodermal, Choroid Plexus, and

Atypical Teratoid – Rhabdoid (ATRT). Such pediatric brain tumors are one of the leading causes of solid tumor cancer – related death in children under the age of 20. Among all childhood brain tumors, about 54% to 70% originate in posterior fossa [3] regions. The posterior fossa is a small space in the skull, found near the brain stem and cerebellum. The cerebellum is the part of the brain responsible for movement. Tumors in this region are considered critical because of limited space within the PF and brain stem nuclei. The PF tumor can block the flow of spinal fluid and cause increased pressure on the brain and spinal cord. Most tumors of the posterior fossa are primary brain cancers, which originate in the brain, rather than spreading from elsewhere in the body. Certain types of PF tumors such as medulloblastoma, ependymomas, primitive neuroectodermal tumors (PNETs) and astrocytoma of the cerebellum and brain stem occur more frequently in children. Some glial tumors such as mixed gliomas are unique in children. They are located in cerebellum (67%) and are usually benign.

Among all PF tumors, brain stem gliomas (BSG) represent 25-30% of all brain tumors, while ependymomas occur 50% in children younger than 3 years [4]. Cystic cerebellar astrocytoma comprises about 33% of all PF tumors in children. It represents 25% of all pediatric tumors. This tumor may be solid or cystic and may be located medially in the vermis or laterally in the cerebellar hemisphere. PNETs include medulloblastomas, medulloepitheliomas, pigmented medulloblastomas, ependymoblastomas, pineoblastomas, and cerebral neuroblastomas. These tumors originate from undifferentiated cells in the subependymal region in the fetal brain. PNETs are second to the cerebellar astrocytoma in frequency, comprising 25% of intracranial tumors in children. These tumors appear heterogeneous on MRI, with areas

of cystic degeneration and central necrosis and isointense in T1 images. Tumors like Ependymomas are derived from ependymal cells. They occur more frequently in females, with 50% presenting in children younger than 3 years. Choroid plexus papilloma and carcinoma represent 0.4-0.6% of all intracranial tumors. They are more frequent in children than in adults (3% of childhood brain tumors). Sixty percent occur in the lateral ventricle and 30% in the fourth ventricle. The third ventricle and cerebellopontine angle are rare locations for this tumor. Dermoid tumors arise from incomplete separation of epithelial ectoderm from neuroectoderm at the region of the anterior neuropore; this usually occurs during the fourth week of gestation. The cyst grows slowly and gradually becomes filled by desquamated epithelium, sweat, and sebaceous materials. More commonly, the cyst occurs in the PF, at or near the midline. Hemangioblastoma represents about 7-12% of all posterior fossa tumors. About 70% of hemangioblastomas occurring in the cerebellum are cystic. Age of presentation is 30-40 years old and is more common in males. Three percent of all cranial metastatic lesions occur in the brainstem and 18% occur in the cerebellum. Originating sites include breast, lung, skin, and kidney. Brainstem gliomas constitute 15% of all brain tumors. In children, brainstem glioma represents 25-30% of all brain tumors. Most brainstem gliomas are low-grade astrocytoma.

The MRI is considered to be the most useful imaging modality for studying brain tissues and tumors. Brain tumor segmentation from MRI is a challenging task because of the heterogeneous appearance in terms of image features such as intensity, texture and shape. The source of heterogeneity is attributed to a) the imaging system and image reconstruction process that is prone to background noise, and b) differentiation interstitial

pressure, perfusion or blood flow over the tumor region. The tumors to be segmented are anatomical structures which are complex in shape, vary greatly in size and position, vary from patients to patients, and may overlap with normal brain tissue. Often a growing tumor can deform the nearby tissues giving it an abnormal appearance. Clinicians and radiologist spend considerable amount of time on segmenting and labeling tumors in MR. There are many works reported in literature [5, 6, 7] on tumor segmentation and identification of tumor region in MR images.

Manual segmentation of tumor tissues is labor intensive. Thus, a computer aided diagnosis (CAD) tool is warranted which can automatically determine the shape, size and volume of the tumor with ease and also reduce the manual efforts saving time. The segmentation task becomes more difficult when one has to drive common decision boundaries on different tissue types in an image. Due to complex structures of different tissues such as white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF) as well as cyst, necrosis and edema in brain, segmentation becomes even more difficult. Consequently, considerable amount of research has been focused on semi-automatic and fully automatic methods for detecting and/or segmenting brain tumors from MRI scans. Zadech et al. [8] has developed an automatic method for adaptive enhancement and unsupervised segmentation of different brain tissues such as CSF, GM and WM in synthetic MR images. Algorri et al. [9] have also used fuzzy parameters to segment normal brain tissues. The most widely performed technique for brain tumor segmentation is using atlas. In atlas based segmentation the template MR volume is registered to the unregistered scans. A one – to –one correspondence between the template and subject images is achieved through high dimensional warping. Warfield et al. [10, 11] combined

elastic atlas segmentation with statistical classification to mask the brain from surrounding structures. Prastawa et al. [12] developed automatic tumor segmentation and statistical classification of brain MR images using an atlas prior. On the computational domain certain machine learning techniques have also been used for MR image segmentation such as self organizing maps (SOM) [13], Support Vector machine (SVM) [14], Markov Random Fields (MRF) [15]. Li et al. [13] has exploited spatial constraints by using a Markov Random Field (MRF) along with self organizing maps model to accurately identify CSF, GM and WM. The MRF takes care of the prior distribution with clique potential and improves the segmentation results without having extra data samples in the training set. Jian et al. [14] has used SVM for separating the tumor and healthy tissues. Gerig et al. [15] have developed a framework in which the information on five properties: voxels intensities, neighborhood coherence, intra-structure properties, inter-structure relationships, and user input flows between the layers via multi-level Markov random fields and Bayesian classification.

Other important technique involved is the feature based technique. Texture can be defined as spatial arrangement of texture primitives or texture elements called textone arranged in more or less periodic manner. Texture primitive, on the other hand, is a group of pixels representing the simple or basic patterns. A texture can be fine, coarse, and smooth or grained depending upon the structure and tone, where tone is based on pixel intensity and structure is the spatial relationship between primitives. The extraction of good features is fundamental to successful image segmentation in this technique. The texture features can capture intensity, irregular variation, mean, variance, skewness, roughness and stochastic process among pixels making it an important. On the other

hand, atlas based techniques, SOM, SVM, MRF cannot capture texture techniques. They can identify the abnormal region by referring a template or by supervised mode.

Texture features have been explored to characterize and segment dystrophic muscles and adipose tissue [16, 17, 18]. Lerski et al. [19] have demonstrated a brain tumor MR image analysis technique, while Mahmoud – Ghoneim et al. [20] have proposed a 3D co- occurrence matrix based tumor texture analysis with increased specificity and sensitivity. However, in both of their works, the volume of interests has been segmented manually. Pachai et al. [21] have proposed a multiresolution pyramid algorithm to segment sclerosis lesions in brain MR image with morphological accuracy and improved reproducibility compared to manual segmentation method. Pitiot et al. [22] have presented a texture based MR image segmentation approach with a novel combination of two – stage hybrid neural classifier. The authors have achieved 90% to 98% classification rate for caudate nucleus, hippocampus and corpus callosum. One of the important features in segmenting tumor from other tissues in brain is intensity. Intensity along with conventional fuzzy c-means clustering algorithm has been used for segmentation of CSF, GM and WM in MR images [23, 24]. However, intensity alone is insufficient to provide successful segmentation. Therefore, other features such as texture type fractal features have been proven effective for analysis of brain tumors. The tumor growth follows fractal process and FD [25] is a natural choice to characterize the textured images and surface roughness. FD has been exploited for to quantify cortical complexity of the brain [26]. Further, multi- fractional Brownian motion (mBm) obtained using stochastic process is shown effective to segment brain tumor [27]. In our previous works [28, 29, 30] the usefulness of intensity, FD and mBm wavelet fractal texture features for

tumor segmentation have been discussed. The tumors are often surrounded by sphere like structures called cyst. A cyst refers to a closed sac that contains fluids, gas or semi solid substances.

1.2. Dissertation Aims

In this dissertation, we investigate computer aided pediatric posterior fossa tumor segmentation. Specifically we investigate efficacy of feature - based tumor segmentation, improvement of segmentation accuracy and validation by a) selecting the best features for Tumor and Non tumor and b) extending the search for best features for abnormalities such as cyst. The goal is to obtain an effective tumor analysis scheme that may ultimately help radiologist and medical physicists in accurate tumor delineations for patient management. Fig. 1.1 shows the proposed outline for achieving our goals.

Consequently, we formulate our three aims as follows. Aim #1 is to select set of features from a given subset of features using an information theoretic approach for PF tumor and non tumors tissues. Aim#2 is to extend the information theoretic approach for multiclass feature selection for selecting the best features for posterior fossa tumor, cyst and non tumor. This aim also obtains improved tumor segments by discriminating cyst from the tumor segments in Aim # 1. Aim # 3 is to obtain an integrated mathematical framework for multiclass feature selection and improved tumor segmentation. Such an integrated mathematical framework will be instrumental in obtaining a complete automated computer aided diagnostic (CAD) system for segmenting PF tumors. We now discuss our aims in more details.

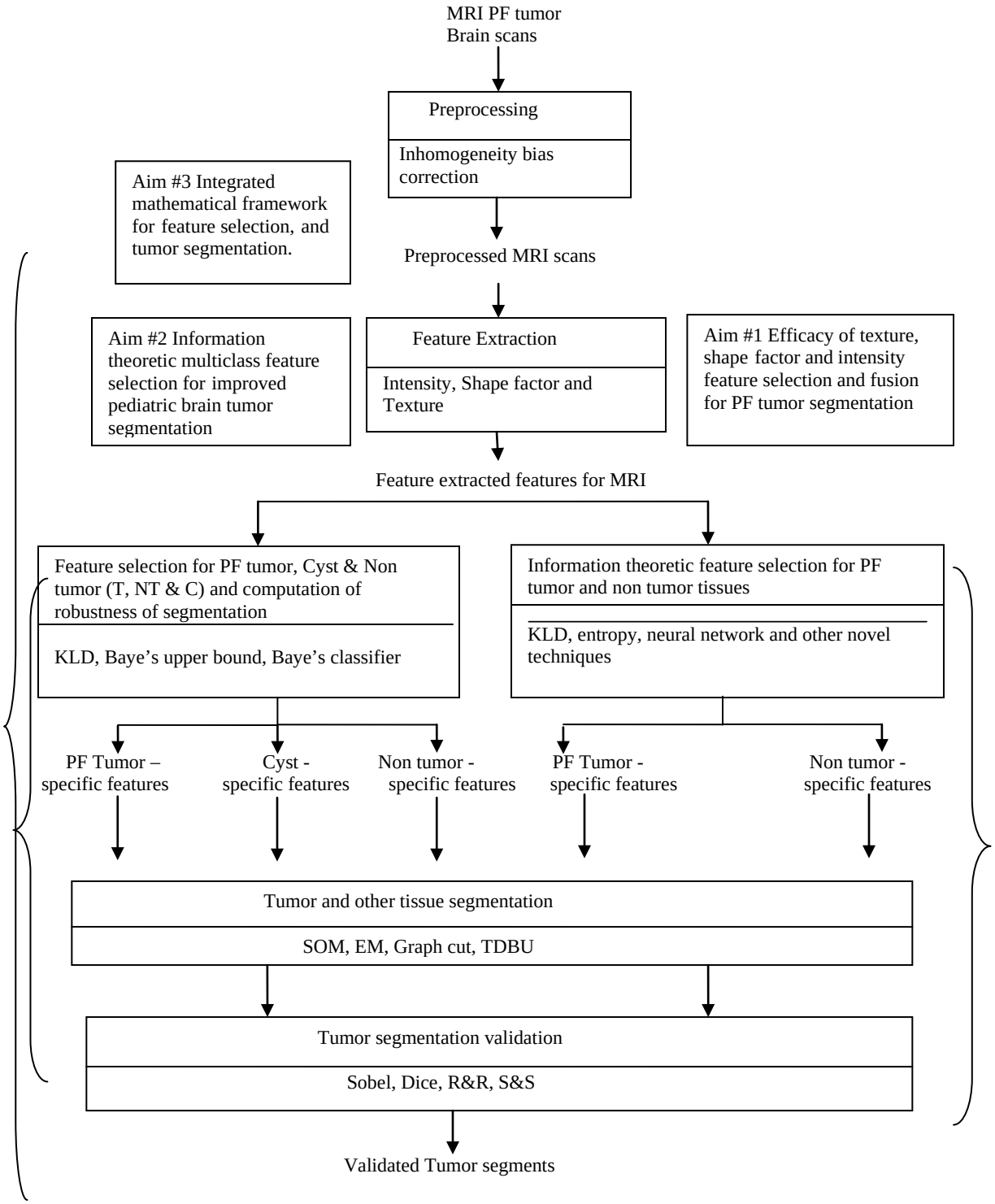


Fig. 1.1 Flowchart showing specific aims

Aim # 1: To investigate the efficacy of texture, shape factor and intensity feature selection and fusion for PF tumor segmentation in MRI.

Rationale # 1

In our previous work [27, 29, 30, 31, 32] we already showed that texture type features such as fractal dimension (FD) and multifractal Brownian motion (mBm) are very useful in segmenting PF tumor from normal brain tissues. However, as we discussed in Section 1.1 feature – based tumor segmentation is an active research areas [16, 17, 18, 19, 20, 21, 22]. Among the different types of features such as intensity, texture, multi-resolution texture and shape factor, some features may be redundant or irrelevant for PF tumor segmentation. Therefore, it is essential to perform systematic feature ranking and selection. Our discussion in Section 1.1 shows that among many different feature selection techniques in literature neural network [33] has been widely used. However, neural network based feature selection is an exhaustive search method; hence, it may be computationally expensive. In comparison, formation theoretic techniques may be more effective for brain tumor feature selection. Among various measures for information theoretic feature selection, we exploit a Kullback Leibler Divergence (KLD) approach for selecting the best features. The KLD is a measure of difference between two probability distributions, whereas other distance measure such as Bhattacharya measures the similarity for two discrete probability classes. Therefore, KLD can be used for multivariate normal distributions, approximated for the class conditional distributions of the tumor and non-tumor regions in MR brain images.

Aim # 2: To investigate information theoretic multiclass feature selection for improved pediatric brain tumor segmentation.

Rationale # 2

The tumors are often surrounded by abnormalities such as edema, cyst and necrosis. In our previous work [27, 29, 30, 31, 32] the segmentation results for tumors included surrounding tissues such as cyst, edema and necrosis. So to improve the accuracy we need to get rid of the surrounding tissues. For achieving this we need specific texture features for surrounding tissues such as cyst. Many works have been done previously for achieving this goal. In Ref. [34] authors have described about the hybrid level set (HLS) segmentation method driven by region and boundary information for segmenting edema and tumor. Region information serves as a propagation force which is robust and boundary information serves as a stopping functional which is accurate. Many neural networks have been exploited for multiclass feature selection. Authors in [35-36] have tried to employ support vector machines (SVMs) to improve the prediction accuracy, and obtained satisfactory results. Further, [37] describes about combination of genetic algorithms with adaboost classifiers to evaluate the effectiveness and the robustness of MNIST database.

On the other hand, many information based criterion has also been explored. Ref. [38] demonstrates the application and impact of the mutual information (MI) criterion for feature selection when developing texture-based CAD tools for the automated diagnostic interpretation of medical images. MI measures the general dependence of random variables without making any assumptions about the nature of their underlying

relationships. Ref. [39] multidimensional local spatial autocorrelation (MLSA) measure that quantifies the spatial autocorrelation of the hyperspectral image data.

The information theoretic approaches described above involve classification of two classes. Since our goal in this dissertation is multiclass separation (i.e. tumor, normal tissues, abnormalities and non tumors) there is a need to investigate multiclass classification algorithm. To achieve this goal we plan to extend KLD to multiclass such that it can be used to select the best features from the given set of features. We also aim at improving the segmentation of tumors by discriminating the cyst.

Aim # 3: To develop an information theoretic mathematical framework for feature selection, and segmentation.

Rationale # 3

Many works has been done for integrating registration, inhomogeneity correction and segmentation of magnetic resonance images. In [40] authors have developed EM framework for estimating the image inhomogeneities, anatomical label maps and mapping from the atlas to image space. The authors in Ref. [41] present an algorithm for adaptive fuzzy segmentation of MRI data and estimation of intensity inhomogeneities using fuzzy logic. Further ref. [42] describes about unifying framework for fully automated inhomogeneity correction and partial volume (PV) segmentation of multi-spectral brain magnetic resonance (MR) images. Warfield et al. [43,44] have combined elastic atlas registration with statistical classification. Elastic registration of a brain atlas helps to mask the brain from surrounding structures.

We are interested in developing an integrated mathematical framework for inhomogeneity correction, optimal feature selection and segmentation of PF tumors in an

EM – KLD framework. In the previous aims feature selection and segmentation are two different methods. In this aim we want to couple feature selection and segmentation. This will help us demonstrate the segmentation at different iterations for feature selection with different features.

1.3. Dissertation Contributions

In this section, the novel contribution of this dissertation is summarized. Note these research contribution follow the above research aims as described above.

1) We investigate the efficacy of several different features, feature ranking and feature selection techniques along with different feature fusion and segmentation methods for PF brain tumor segmentation using the selected features. The novelty comes from the fact that we implement an integrated mathematical framework for features selection using Kullback Leibler Divergence for PF tumors in pediatric brain MRI.
2) To improve the tumor segmentation we extend the two class KLD techniques to multiclass feature selection techniques, so that we can effectively select features for tumor, cyst and non tumor tissues. We further obtain segmentation robustness for each tissue types by computing the Bayes posterior probabilities and corresponding number of pixels for each tissue segments in MRI patient images. For KLD computes the differences between the conditional probabilities for two classes. Bayes upper bound property minimizes the error of classifier by selecting the features taking into account their effects on classification errors. The novelty comes from the approach that we combine these properties of KLD and Bayes

upper bound to develop a multiclass feature selection for tumor, cyst and non tumor tissues.

3) A novel integrated information theoretic mathematical framework for inhomogeneity correction, feature selection using KLD and tumor segmentation in EM is developed for pediatric PF tumors. To the knowledge, no integrated framework has yet been proposed to combine information from intensity inhomogeneity to segmentation for automated segmentation of PF tumors.

1.4. Dissertation Organization

The remainder of the dissertation is organized as follows. In Chapter 2, a detailed background review for all the 3 aims proposed has been described. Chapter 3 describes about the "Efficacy of texture, shape and intensity feature fusion for PF tumor segmentation in MRI". Chapter 4 mentions about the "Information theoretic multiclass feature selection and improved pediatric brain tumor segmentation robustness evaluation". In Chapter 5 describes "Develop an information theoretic mathematical framework for feature selection, and segmentation". Introduction, methods, results and conclusion have been described in Chapter 3, 4 and 5. Finally chapter 6 provides the concluding remarks and some future directions.

2. Literature Review

In this section, we first review relevant background for feature extraction using fractal, multifractal texture and level set shape methods. The tumor growth is known to follow a fractal process [45] that can be quantified using FD. The FD can be used as a measure of degree of the texture complexity of the tumor surface. Among many other conventional feature extraction methods, the Gabor filters are suitable to capture discontinuity in intensity and texture in an image [46]. Wavelet – Gabor filters have been investigated to outline the area of metaplastic changes in cervical images [47], and also to differentiate prostate and non- prostate tissues [48]. However, the wavelet – Gabor technique does not provide an integrated mathematical framework for simultaneous analysis of tumor texture at different resolutions. In comparison, FD is inherently suitable to capture salient fractal properties in an image such as self-similar features in addition to texture variation. In addition, wavelet-fractal techniques capture the multiresolution and texture features simultaneously for effective tumor segmentation [27, 28, 29, 49].

Among many different types of features such as intensity, texture, multi-resolution texture and shape, some features may be redundant or irrelevant for PF tumor segmentation. Therefore, it is essential to perform systematic feature ranking and selection. Among many different feature selection techniques neural networks [50] has attracted attention. However, neural network based feature selection is an exhaustive search method; hence it may be computationally expensive. Another hybrid technique that uses classifiers is known as Boosting [51]. In addition, simple techniques such as PCA [52] have also been used for feature selection. On the other hand, the KLD provides a quantitative feature ranking considering the entropy gain of features and ranks them is a

decreasing order. In this work, we systematically investigate four different feature selection techniques such as use Kullback – Leibler Divergence (KLD) [53], Principal Component Analysis (PCA) [52], Boosting [51] and entropy to ascertain which features offer the maximum separability between PF tumor and non- tumor tissues. For feature fusion and segmentation, there have been various methods reported in literature [54, 55]. In Ref. [54], the authors have proposed novel idea of combining Top Down and Bottom Up (TDBU) segmentation. Among different feature fusion and clustering tools [56] EM algorithm is an efficient iterative procedure to compute the Maximum Likelihood (ML) estimate in presence of missing or hidden data [56]. Note that KLD and EM can be combined in a single mathematical framework for feature selection and segmentation. In this work, we explore effectiveness of three different feature fusion and segmentation techniques such as EM, TDBU and graph cut respectively.

2.1. Fractal Dimension (FD) texture feature extraction

The concept of fractal is first proposed by Mandelbrot [57] to describe the geometry of the objects in nature. The FD is a real number that characterizes the fractalness (texture) of the objects. We investigate effectiveness of three different FD computation methods for brain tumor segmentation in MRI [31]. In a prior work [31], we demonstrate that piecewise- triangular-prism-surface-area (PTPSA) method offers the most reliable FD values and resulting tumor segmentation.

2.2. Multifractional Brownian Motion (mBm) texture feature extraction

We have successfully investigated mBm-based texture model for brain tumor segmentation in MRI [27]. The mBm is defined as,

$$x(at) = a^{H(t)} x(t), \tag{1}$$

where x(t) is an mBm process, H(t) is the time varying scaling (or Holder) exponent and a is the scaling factor. After a sequence of mathematical derivation, the expectation of the squared –magnitude of the wavelet transform, W_x, of x(t) is given as,

$$\log\left(E\left[\left|W_x(t,a)\right|^2\right]\right) = (2H(t)+1)\log a + CONSTANT, \tag{2}$$

The critical research issue is to obtain a robust estimation of the expectation of the squared-magnitude of the wavelet coefficients given a single observation of the random process x. Following Goncalves et al. [58], the empirical estimate of the q-th order moment of $\left|W_x(t,a)\right|$ can be estimated as,

$$E\left\{\left|W_x(t,a)\right|^q\right\} = \frac{1}{N}\sum_{i=0}^{N-1}\left|W_x(t_i,a)\right|^q, \tag{3}$$

where single realization of the analyzed process is sampled on a uniform lattice $t_i = i/N$, i = 0, …., N-1). Plugging q = 2 in Eq. (4) and combining with Eq. (3) one can approximate H(t) as follows,

$$2H(t) = \lim_{a \to 0^+} \frac{\log\left(\frac{1}{N} \sum_{i=0}^{N-1} \left| W_x(t_i, a) \right|^2 \right)}{\log a}. \tag{4}$$

For 2-D mBm model, let $z(\vec{u})$ represent a 2D mBm process, where $\vec{u}$ denotes the vector position (u_x, u_y) of a point in the process. Following the similar derivations for Eqn. (2), we approximate H ($\vec{u}$) for a 2D mBm process as follows [27],

$$2H(\vec{u}) = \lim_{a \to 0} \frac{\log\left(\frac{1}{M+N} \sum_{x=o}^{N-1} \sum_{y=0}^{M-1} \left| W_z(b_{x,y}, a) \right|^2 \right)}{\log a}. \tag{5}$$

The Eqn. that links H ($\vec{u}$) with FD is given as,

$$FD = E+1-H. \tag{6}$$

where E is Euclidean Dimension of the space of fractal (E = 2 for 2D image) and H is the Hurst coefficient.

2.3. Level set based shape feature extraction

Level-set based shape modeling is an important research topic in computer vision and computer graphics. Shape models aid the tasks of object representation and recognition. The authors in [59] modified the level set method first proposed in [60] and obtain a model for shape-based representation of objects in image. In this study, we implement a more recent work [61] on binary level set representation for object shape detection. Consider the basic definition of level given as [60],

$$\phi_t + F|\nabla\phi| = 0, given \phi(x, t = 0), \tag{7}$$

and

$$\phi_t + F_o|\nabla\phi| + U(x, y, t)\nabla\phi = \varepsilon K|\nabla\phi| \ . \tag{8}$$

where $F_o|\nabla\phi|$ is the motion of curve in the direction normal to front, $U(x, y, t)\nabla\phi$ is the term that moves the curve across the surface, $\varepsilon K|\nabla\phi|$ is the speed term dependent upon

curvarture. In our work, $U(x, y, t)$ is the gradient of image and $\varepsilon K|\nabla\phi|$ is approximated using a central difference. We first convert the MRI to binary image. The level set is used on these binary images to track the shape at the boundary of images. Note, for binary images, only digital derivative approximations exists at the boundary. We initialize the level set function using the gradient of the image. We propagate this gradient across the surface given as [61],

$$\phi_{ij}^{t+1} = \phi_{ij}^{t} - \Delta t[\max(G_{ij},0)\Delta^{+} + \min(G_{ij},0)\Delta^{-}]. \tag{9}$$

where ϕ_{ij}^{t} is the value of ϕ at pixel *i* at time *t*, Δt is the time step (or scaling factor), G_{ij} is a Gaussian filter to smooth the edges and Δ^{+} and Δ^{-} describe the normal component and given as,

$$\Delta^{+} = [\max(D_x^{-},0)^2 + \min(D_x^{+},0)^2 + \max(D_y^{-},0)^2 + \min(D_y^{+}0)^2]^{1/2}, \tag{10}$$

and

$$\Delta^{-} = [\max(D_x^{+},0)^2 + \min(D_x^{-},0)^2 + \max(D_y^{+},0)^2 + \min(D_y^{-},0)^2]^{1/2}. \tag{11}$$

where $D_x^-, D_x^+, D_y^-, D_y^+$ are the forward and backward derivative approximation in x and y directions. These steps iterate and stop when the boundary is completed upon convergence.

2.4. Kullback – Leibler Divergence for Feature selection & Entropy for feature ranking

The KLD is a measure of difference between two probability distributions [53]. Therefore, KLD can be used for multivariate normal distributions, approximated for the class conditional distributions of the tumor and non-tumor regions in MR brain images. The equation for the parametric model for ω-th class conditional density function for a random variable x is given as [53],

$$p(x|\omega|) = \sum_{m=1}^{M\omega} \alpha_m^{\omega} g_o(x|\mu_o, \sigma_o) g(x|\mu_m^{\omega}, \sigma_m^{\omega}, \mu_o, \sigma_o, \theta), \tag{12}$$

where,

$$g_o(x|\mu_o, \sigma_o) = \prod_{i=1}^{D} \left[\frac{1}{\sqrt{2\pi}\sigma_{0i}} \exp\left\{ \frac{-1}{2} \left(\frac{x_i - \mu_{0i}}{\sigma_{0i}} \right)^2 \right] \right], \tag{13}$$

and,

$$g_o(x|\mu_m^{\omega},\sigma_m^{\omega},\mu_o,\sigma_o,\theta)=\prod_{i=1}^{D}\left[\frac{\sigma_{oi}}{\sigma_{mi}^{\omega}}\exp\left\{\frac{-1}{2}\left(\frac{x_i-\mu_{mi}^{\omega}}{\sigma_{mi}^{\omega}}\right)+\frac{1}{2}\left(\frac{x_i-\mu_{oi}}{\sigma_{oi}}\right)^2\right\}\right]^{\theta i} . \tag{14}$$

(μ_o, σ_o) is the mean and variance for first class, (μ_m, σ_m) is the mean and variance for the second class and θ is the control parmeter. α_m^{ω} is a non negative weight, $\sum_{m=1}^{M_{\omega}}\alpha_m^{\omega}=1$ and M_{ω} is the number of features component. We now consider the maximum likelihood estimation (MLE) of all the unknown parameters such as $A=A_{\omega}, B=B_{\omega}, \mu_o$ and σ_0 in the parametric family. We use EM algorithm to maximize the log likelihood function w. r. t. parameters A, B, μ_o, σ_o with given θ. The KLD between two classes ω_1 and ω_2 is given as [53],

$$J_i^{\omega}=\frac{1}{2}\left\{\sum_{m=1}^{M_{\Omega}}\alpha_m^{\omega}\log\left(\frac{1}{\sigma_{mi}^{\omega}}\right)^2-1+\sum_{m=1}^{M_{\Omega-\omega}}\alpha_m^{\omega,\Omega-\omega}\left[\log\left(\sigma_{mi}^{\Omega-\omega}\right)^2+\left(\frac{\sigma_{mi}^{\Omega-\omega}}{\sigma_{mi}^{\omega,\Omega-\omega}}\right)^2+\left(\frac{\mu_{mi}^{\omega,\Omega-\omega}-\mu_{mi}^{\Omega-\omega}}{\sigma_{mi}^{\Omega-\omega}}\right)^2\right]\right\} . \tag{15}$$

where,

$$\mu_{mi}^{\omega}=\sum_{x\in X\omega}x_i g\left(x|\mu_0,\sigma_0,\omega\right); \tag{16}$$

$$\mu_{mi}^{\omega,\Omega-\omega} = \sum_{x \in X\omega} x_i g^{\omega}\left(x \middle| \mu_0, \sigma_0, \Omega - \omega\right); \tag{17}$$

$$\left(\sigma_{mi}^{\omega}\right)^2 = \sum_{x \in X\omega} \left(x_i - \mu_{mi}^{\omega}\right)^2 g\left(x \middle| \mu_0, \sigma_0, \omega\right); \tag{18}$$

$$\left(\sigma_{mi}^{\omega,\Omega-\omega}\right)^2 = \sum_{x \in X\omega} \left(x_i - \mu_{mi}^{\omega,\Omega-\omega}\right) g^{\omega}\left(x \middle| \mu_0, \sigma_0, \Omega - \omega\right). \tag{19}$$

and

$$J_i = \sum_{i=1}^{D} \theta_i J_i^{\omega}. \tag{20}$$

Figure 2.1 shows the formal algorithm for KLD computation.

KL divergence (X, N, k. Ji)

X is the input matrix of size n x1. N is the number of features/dimensions. K is the desired number of clusters.

1. Compute the weights $g_0(x|\mu_0, \sigma_0)$ *and* $g_0(x|\mu_m, \sigma_m)$ *using Eqns. (13) and (14)*

2. Under fixed weights compute the value of μ_{mi}^{ω}, $(\sigma_{mi}^{\omega})^2$, $\mu_{mi}^{\omega,\Omega-\omega}$, $(\sigma_{mi}^{\omega,\Omega-\omega})^2$ *using Eqns.(16)-(19)*

3. Using the parameters of μ_{mi}^{ω}, $(\sigma_{mi}^{\omega})^2$, $\mu_{mi}^{\omega,\Omega-\omega}$, $(\sigma_{mi}^{\omega,\Omega-\omega})^2$ *and weights compute the value of the KLD using Eqn. (15)*

4. Compute the entropy using Eqn.(21)

Fig. 2.1 Algorithm for computing the feature selection and ranking [53]

We exploit the idea of information theory such as mutual information and KLD for feature ranking and selection. The mutual information can also be understood as the expectation of the KLD of the univariate distribution $p(x)$ of X from the conditional distribution $p(x|c)$ of X given C. This suggest that the more different the distributions $p(x|c)$ and $p(x)$, the greater is the information gain, I(x,c) as follows,

$$\text{Gain} = I(x,c) = E_{xp}\{D_{KL}(p(x|c)\|p(x)\} \quad (21)$$

According to feature ranking based on information, gain ranks feature X over feature Y if Gain (X, C) > Gain (Y, C). Therefore, a feature should be ranked if it can reduce more entropy than the other. We find the entropy for all the four features – intensity, mBm, fractal and shape using Eqn. (21).

2.5. Similarity Coefficient (SC) for segmentation quality and robustness identification

For estimating the robustness of segmentation we consider different similarity measures such as Jaccard, Dice, Sokal & Sneath (SS) and Russel and Rao (RR) [29]. Note the study of outlier and its effect on segmentation and pattern classification is better understood using region of curves (ROC) analysis, which is beyond the scope of this work. We quantify segmentation robustness by measuring the overlap of tumor using different similarity metrics such as Jaccard ($p/p+q$), Dice ($2p/2p+q$), SS $(p/p+2r)$ and RR $(p/p+q+r)$, where p is the area of tumor region in MRI (tumors segmented by radiologist and used as ground truth), q is the area of the tumor region segmented using EM algorithm and r is the non-tumor region. Note computations of both Dice and Jaccard involve the ratio between actual and automated tumor segments. On the other hand, SS and RR involve computations of both the ratio between actual tumor segments to automated tumor segments and the non-tumor regions. Overall, these overlap ratios indicate the accuracy of tumor segmentation results for each patient. Value of 1.0 for any of these measures represents complete overlap whereas 0.0 represents no overlap.

In this section, we first review relevant background for different feature selection techniques. In machine learning and statistics, feature selection, is the technique of selecting a subset of relevant features for building robust learning models. Two class feature selection has many limitations. Methods such as classifier LDA include the fact that the classifier must have a linear form. The performance degrades when the two groups to be classified are not perfectly separable in feature space. For very large training set, the minimum error rate in feature space is not achievable. These problems are

especially relevant in medical image segmentation because of the large image dataset and the high complexity of the images. Wrapper-based feature selection is attractive because wrapper methods are able to optimize the features they select to the specific learning algorithm. Unfortunately, wrapper methods are prohibitively expensive to use with neural nets.

There have been works in information theoretic methods for feature selection. In Ref. [62], separability indices, such as Bhattacharyya distance, Jeffries – Matusita distance and Mahalanobis distance have been used to determine the best band combination in multitemporal dataset. A probabilistic Bayesian network model is used in Ref. [63] to systematically select the representative performance features, which can provide optimal classification accuracy and adapt to changing workloads. In Ref. [64], a novel feature selection scheme based on the upper bound of Baye's error under normal distribution for the multi-class dimension reduction problem is proposed. In order to obtain an accurate solution of the feature selection transform matrix in term of the minimum upper bound of Baye's error, a recursive algorithm based on gradient method is developed.

Bruazzone et al. [65-68] has extended two class classifications to multiclass for satellite images. The authors in Refs. [65, 66] discuss multiclass features selection for distance measure such as Bhattacharya, Jeffery - Matsutia, F index and Baye's criterion for remote sensing images. Reference [67] discusses a data fusion approach to the classification of multisource and multitemporal remote-sensing images. The method is based on the application of the Bayes rule for minimum error to the "compound" classification of pairs of multisource images acquired at two different dates. In particular,

the fusion of multisource data is obtained by using multilayer perceptron neural networks for a nonparametric estimation of posterior class probabilities. Furthermore, a supervised nonparametric technique based on the “compound classification rule” for minimum error is discussed in [68] to detect land-cover transitions between two remote-sensing images acquired at different times. The methods in Ref. [68] offer discriminative feature subset as a group rather than emphasizing the individual contribution of features to discriminate. In addition, the probabilistic distance measures which are used as criterion function offers the individual discriminatory powers of features.

In this work, we propose to improve our prior tumor segmentation accuracy by segmenting cyst tissues from tumor regions using information theoretic KLD method. In our prior works [69, 70] the segmented tumor regions include other non-tumor tissues such as cyst, edema and necrosis. In this work, we develop an integrated probabilistic KLD technique for multiclass feature selection and improved pediatric brain tumor segmentation. KLD computes the difference between the conditional probabilities for two classes. The greater the difference the best is the separation between classes. Baye’s upper bound property minimizes the error of classifier by selecting the features taking into account their effects on classification errors [65]. We combine these properties of KLD and Baye’s upper bound to develop a multiclass feature selection for tumor (T), cyst (C) and non tumor (NT) tissues. We further evaluate our improved tumor segmentation robustness using evaluation metrics for eight patients in T1, T2 and FLAIR modalities. The criterion we present in this paper is based on upper bound to the Baye’s error formulated under appropriate simplifying hypotheses. We define the criterion for two-class case and then generalize it to multiclass feature selection.

3.1. Multiclass feature selection Criteria

Several distance measures such as Bhattacharya and Jeffrey's – Matusita (JM) have been used for statistical separability of classes. The Bhattacharya distance measure is given as, [65],

$$B_{ij} = \ln\left\{\int_x \sqrt{p(x|\omega_i)\,p(x|\omega_j)}\,dx\right\}, \tag{22}$$

where $p(x|\omega_i)$ and $p(x|\omega_j)$ are the conditional probabilities for two classes. The JM distance, on the other hand, is given as, [65],

$$J_{ij} = \left\{\int_x \left[\sqrt{p(x|\omega_i)} - \sqrt{p(x|\omega_j)}\right]^2 dx\right\}^{1/2}. \tag{23}$$

Consider two classes, ω_i and ω_j. The error probability of the Baye's classifier for the minimum error is given as, [65],

$$P_e(\omega_i,\omega_j) = P(\omega_i)\int_{x\in D_j} P(\xi|\omega_i)d\xi + P(\omega_j)\int_{x\in D_i} P(\xi|\omega_j)d\xi \quad , \tag{24}$$

where D_i and D_j are the decision regions in the feature space X for the classes ω_i and ω_j, respectively. The D's are defined as,

$$D_i = \left\{ x \in X \middle| P(\omega_i)\, p(x|\omega_i) \right\} \geq P(\omega_j)\, p(x|\omega_j) \quad , \tag{25}$$

and

$$D_j = \left\{ x \in X \middle| P(\omega_j)\, p(x|\omega_j) \right\} \geq P(\omega_i)\, p(x|\omega_i) \ . \tag{26}$$

Assuming Gaussian distributions, Eqn. (24) can be written as,

$$P_e(\omega_i, \omega_j) = P(\omega_i)\left[1 - Q\left(\frac{\alpha - 1/2d_{ij}}{\sqrt{d_{ij}}}\right)\right] + P(\omega_j) Q\left(\frac{\alpha + 1/2d_{ij}}{\sqrt{d_{ij}}}\right) \quad , \tag{27}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\xi^2/2} d\xi$. The value of α depends on the optimal decision parameter which is computed using maximum a posterior probability (MAP) as follows,

$$\alpha = \log P(\omega_i) / P(\omega_j). \tag{28}$$

The priors are computed as the number of pixels covering the tumor, cyst and non tumor region to the total number of pixels.

3.2. Segmentation using mBm feature

For effective segmentation, we consider each pixel within a region share similar characteristic such as intensity or texture with its neighbors. To be specific, a pixel within a texture region can be considered as an interior pixel for a neighborhood in the raw image. On the other hand, boundary pixel does not share similar property with its neighbors. This allows effective separation of interior and boundary pixels for a texture region.

Once we select the best feature using Bayesian KLD technique, we obtain segmentation robustness for selected feature at pixel level. We compute the number of pixels for segmented tumors using the best feature. Texture feature such as FD and mBm are non linear feature extraction process. Therefore, in order to compute pixel level accuracy for tumor segments we consider sub images which cover the tumor region, and then obtain a suitable threshold for interior pixels and exterior pixels for selecting from those sub images.

Let us consider c (i, j) to represent the number of sub images containing pixel x (i, j). Let also x (i, j) be an interior point of a certain texture region and CM (i, j) represents the regions in the sub images. Assume that the number of sub images used in the algorithm is K; the possible value of interior threshold (Th_{int}) ranges from 1 to K. To determine the best values of Th_{int}, a two-step method is adopted [71]. The steps involved to testing each candidate value for Th_{int} and determination of the best threshold are based

on number of regions created as well as the separability among these regions. Considering all possible threshold candidates, the most frequent resulting region number is determined first. The threshold resulting from the corresponding region number is chosen as the tentative candidate for Th_{int}. If there are more than one most frequent resulting region numbers, then the threshold that results in larger class separability is selected as the final threshold. In order to obtain an automatic threshold selection algorithm, let us define a few terms.

The class separability (SP) is defined as the ratio of intra-distance, d_{intra} and inter-distance, d_{inter} of the texture regions, TR_j, j = 1, 2,….., m. SP = d_2/d_1 . Let M_j denote the mean gray value of TR_j, and $p_j(x, y)$ be the value of the pixel (x, y) in TR_j. Then, d_{intra} and d_{inter} are defined as follows,

$$d_{intra} = \frac{1}{m}\sum_{j=1}^{m}\frac{1}{A_j}\sum_{(x,y)\in TRj}\left(p_j(x,y) - M_j\right)^2 ; \tag{29}$$

and,

$$d_{inter} = \frac{2}{m(m-1)}\sum_{k=1}^{m}\sum_{\substack{j=1,\\ j\neq k}}^{m}\frac{1}{A_j}\sum_{(x,y)\in TR_j}\left(p_j(x,y) - M_k\right)^2 ; \tag{30}$$

where A_j is the area size of TR_j. The algorithm for computing the interior pixels is given below;

Algorithm for computing interior pixels:

1) For each sub image s_k covering T or C region do

a) For each pixel x (i, j) in s_k do

Check the M XM neighborhood N (i, j) of x (i, j)

If more than p% pixels of N (i, j) are in the same region as x (i, j) then

c (i, j) = c (i, j) + 1

2) For each counter c (i; j)

b) Perform a region growing to produce texture region TR_{ij} = 1,2,….m

c) Compute separability for different regions SP using eqn. (10)

d) Select SP_{Max} = 0.

e) For each texture region TR if

If m_i = m and

If SP_i >MaxSP then

MaxSP = SP_i.

Th_{int} = Threshold Candidate.

Fig.2.2 Algorithm for computing the interior pixels for texture (mBm) feature [71].

We separate the interior of texture regions from their boundaries using algorithms in Fig.2.2. The texture regions, TR_{ij}, are obtained by region growing. Given a boundary pixel *p* and a sub image *s*, we locate the texture region r_p in *s* in which *p* resides, then determine the texture region in TR_{ij} which overlaps most with r_p. After checking all sub images, *p* is assigned to the texture region in TR_{ij} to which it is assigned most frequently. We show boundary pixel assignment algorithm as follows.

Algorithm: Boundary Pixel Assignment

1) For each boundary pixel p do

a) *For each sub image s_k do*

Determine the region, r_k in s_k, to which p belongs based on the clusters of s_k

Determine the region r_m of TR_{ij}, which overlaps with r_k most

$cm = cm + 1$

End

$j = argmax_i\ c_i$

Assign p to region r_j

End

Fig. 2.3 Algorithm for computing the boundary pixels for texture (mBm) feature [71].

In this work we review relevant work based on statistical approach for segmentation of structures using inhomogeneity. An elaborate initialization scheme is suggested to link the set of Gaussians per tissue type, such that each Gaussian in the set has similar intensity characteristics with minimal overlapping spatial supports [72]. Segmentation of the brain image is achieved by the affiliation of each voxel to the component of the model that maximized the a posteriori probability. Incorporating spatial information via a statistical atlas provides a means for improving the segmentation results [73, 74, 75]. The statistical atlas provides the prior probability for each pixel to originate from a particular tissue class. Algorithms that are based on the maximum a posteriori (MAP) criterion utilize the atlas information in the algorithm iterations to augment the information in the presence of noisy data. In Ref. [76] a three-step segmentation procedure is discussed. First, segmentation of brain/non-brain tissue is performed by using Hybrid watershed algorithm (HWA). Then the intensity inhomogeneity correction method is applied to MR image. Finally, Fuzzy Kohonen's Competitive Learning (F-KCL) Algorithms are used for MRI tissue segmentation.

Recently EM approaches have been utilized for computing the bias field. Wells et al. [77, 78, 79] propose an expectation-maximization (EM) algorithm to achieve an interleaved bias correction/statistical segmentation. In the case of scalar data, the bias estimate $\hat{b}$ is calculated as $\hat{b} = H[Y - WU]$, where H is a low-pass filter, Y the original data and WU a prediction of the signal, which is the sum of the class means weighted by the a posteriori probabilities, $\sum P_c \mu_c$. Wells' formulation includes the bias distortion in the statistical model of the pixel distribution, i.e. the bias field influences the distribution by locally shifting its mean value. The algorithm iterates between two steps, the E-step

for calculating the posterior tissue probabilities, and the M-step for estimating the bias field. Regis et al. [80] introduced a modified EM algorithm that replaces the distribution of the class other, which includes all tissue not explicitly modeled, by a uniform probability density function. The correction claims to be more robust and to overcome some limitations of Wells' original method. They also introduce an automatic estimation of the initial parameters based on a constrained and exhaustive search guided by minimum entropy. Nevertheless, the initialization of the parameters remains critical, as in the original algorithm, and the method is still sensitive to the spatial configuration of image structures. Leemput et al. [81, 82] developed fully automatic segmentation of MR head images by statistical classification using an atlas prior both for initialization of probability density functions and also for geometric constraints, solved as an expectation maximization (EM) algorithm. The method has been shown to be very robust and highly reproducible for normal brain images, but fails in the presence of large pathology.

In this work, we have not considered registration as a parameter. Registration of brain MRI having tumor if registered on atlas will produce erroneous result. The brain tumors can't be simply modeled as intensity outliers due to overlapping intensities with normal tissue and/or significant size. In [83], for example, a criterion for detecting outliers is used to generate a tumor prior in a subsequent EM segmentation which is treating tumor as an additional tissue class. Ref. [84] introduces a generative probabilistic model for segmentation of tumors in multi-dimensional images. The model allows for different tumor boundaries in each channel, reflecting difference in tumor appearance across modalities. They augment a probabilistic atlas of healthy tissue priors with a latent

atlas of the lesion and derive the estimation algorithm to extract tumor boundaries and the latent atlas from the image data.

So far there has been work on combining registration and inhomgeneity to see the effect of segmentation [85]. But in this work, we are trying to see the effect of segmentation when varying textural features. This is achieved in a single framework by combining Inhomogeneity, feature selection and segmentation. In this work we combine feature selection method with homogeneity in an Expectation Maximization framework to study that feature selection also plays an important role for improving tumor segmentation.

3. Efficacy of Texture, Shape and Intensity Feature Fusion for Posterior - Fossa Tumor Segmentation in MRI

3.1. Introduction

Brain tissue and tumor segmentation in MR images has been an active research area. Extraction of good features is fundamental to successful image segmentation. Due to complex structures of different tissues such as the gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) in the MR brain images, extraction of useful features is a challenging task. Variability in tumor location, shape, size and texture properties further complicates the search for robust features. Posterior fossa (PF) tumor is usually located near the brain stem and cerebellum. About 55%- 70% pediatric brain tumors arise in the PF, compared with only 15%- 20% of adult tumors. Most tumors of the PF are primary brain cancers, which originate in the brain, rather than spreading from elsewhere in the body. Due to narrow confinement at the base of the skull, complete removal of PF tumors poses non-trivial challenges. Therefore, accurate segmentation of PF tumor is necessary.

Intensity is an important feature in segmenting tumor from other tissues in the brain. In Ref. [86], the authors use intensity and a conventional fuzzy c-means clustering algorithm for segmentation of CSF, GM and WM in MR images. However, using intensity alone for segmentation has proved to be insufficient. Fractal Dimension (FD) is a useful tool to characterize the textured images and surface roughness [87]. FD has been exploited to quantify cortical complexity of the brain [88]. Further, texture feature obtained using a stochastic Multi- fractional Brownian motion (mBm) model is shown effective to segment brain tumor [28]. In our previous works [28, 29, 30, 31], we discuss

the usefulness of intensity, FD and mBm wavelet fractal texture features for tumor segmentation. However, for patients with poor MR image quality, the texture and intensity features may prove inadequate for PF tumor segmentation. For these patients, another feature such as the shape may be useful for improved PF tumor segmentation in MR images.

The level set method first developed by Osher et al. [89] has found applications in many disciplines such as image processing, computer graphics, computational geometry and optimization. The level set is a numerical analysis technique for tracking interferences and shape. Some applications of level sets in medical image analysis are extraction of complex shapes such as the human cortex in MRI for neurological disease diagnosis [60] and shape-based approach to curve evolution for the segmentation of medical images [90]. In a recent work [61], a binary level set method has been introduced to reduce the expensive computational cost of redistancing the traditional level set function.

Feature selection, also known as variable selection, feature reduction, attribute selection or variable subset selection, is a technique for selecting a subset of relevant features for building robust learning models. Feature selection has been exploited in many applications such as medical imaging, data mining and lexical works [91, 92]. In medical imaging, various techniques have been used to select the best features from a given set of features [93, 53]. A new feature selection technique based on KLD between two-class conditional densities functions approximated by finite mixture of parameterized densities has also been discussed [53].

In our current work, we evaluate the efficacy of the level set shape along with fractal texture and intensity features to discriminate PF tumor from other tissues in the brain. We investigate the efficacy of several different features, feature ranking and feature selection techniques along with different feature fusion and segmentation methods for PF brain tumor segmentation using the selected features.

3.2. Methods and Datasets

The overall flow diagram our method is shown in Fig.3.1. The first step includes the preprocessing stage that minimizes this intensity bias using the normalization algorithm. After intensity normalization we extract four features such as intensity, FD using PTPSA algorithm, mBm using fractal- wavelet algorithm and shape using level set method in multimodality MR images. We use both KLD and the entropy values for feature ranking and selection. The features selected are then used for the segmentation of the tumor region in MRI using EM. Note, an integrated EM framework allows us to obtain feature ranking and selection using KLD and subsequent tumor segmentation simultaneously.

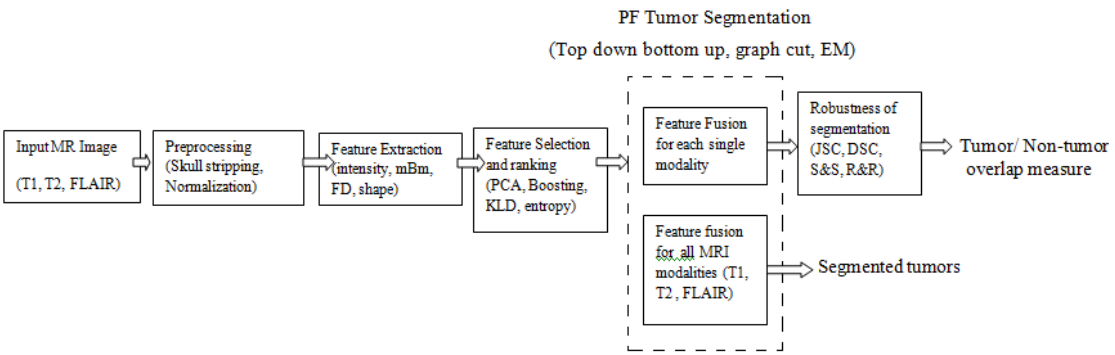


Fig. 3.1 Flow diagram showing the steps

3.2.1. Image Intensity Normalization

The MRI intensity is affected by various sources of variations such as different parameter settings and physics of imaging device. To minimize the intensity bias of the MR image, intensity normalization is used as pre – processing step. In this work, we implement a two- step normalization method in [27, 29] where the image histograms are modified such that the histograms match a mean histogram obtained using the training data. After applying the normalization method, the intensity values for the same tissue in different MR images fall into a very narrow range (usually a single value) in the normalized image.

3.2.2. Feature Extraction

After intensity normalization and bias correction we extract three sets of features from the normalized images in T1, T2 and Flair modality.

3.2.3. Feature selection and feature ranking using different methods

The authors in [53] propose a novel PCA – based method for dimensionality reduction of features known as Principal Feature Analysis (PFA). The PFA has been successfully applied for choosing the principal features in face tracking and content based image retrieval problems. Similarly, a boost feature subset selection (BFSS) method has been proposed to select and rank genes in microarray data on the basis of discriminative scores to improve the performance [52]. For BFSS implementation, we compute the F-score for each feature type in each modality. We then rank the F-scores in the descending order. For comparison, we also formalize the mathematical derivation of KLD. We exploit KLD for ranking and selecting the best feature combinations among four features for tumor / non-tumor discrimination. After the feature has been extracted, we use the KLD algorithm in Fig. 2.1 for feature selection.

3.2.4. Image Segmentation using different algorithms

We study three different segmentation techniques for comparison. We first implement TDBU method as follows. For top-down step, we extract the texture features from MRI slices and obtain boundary of the tumors and non-tumor region based on the mutual information. Then to avoid inconsistent boundary in the top down step, we perform bottom up processing which segments the tumor by considering the coherent groups of pixels that belong to tumor based on the texture features. For graph cut method [94], the image is considered a graph and nodes i and j are pixels. Note the edge weight W_{ij} denotes a local measure of similarity between two pixels. Let G = {V, E} where V

stands for the node and E for edges. The similarity between two groups is called a cut and is given by,

$$cut(A,B) = \sum_{i \in A, j \in B} W_{ij} \tag{31}$$

The fundamental issue is specifying the edge weights W_{ij} for which we rely on the normalized cuts. Shi et.al [95] proposed a normalized cut to separating the region and defined normalized cut as follows,

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(B,A)}{assoc(B,V)}, \tag{32}$$

where $assoc(A,V) = \sum_{i \in A, k \in V} W_{ik}$ is the total connection from node A to all the nodes in the graph. In our work, we compute the edge weights W_{ij} between the two pixels as follows,

$$W_{ij}^{TX} = \exp\left(\frac{\left\| FD_i - FD_j \right\|}{FD_i + FD_j} + \frac{\left\| mBm_i - mBm_j \right\|}{mBm_i + mBm_j} + \frac{\left\| shape_i - shape_j \right\|}{shape_i + shape_j}\right). \tag{33}$$

where, $W_{ij} = W_{ij}^{TX}$ .We compute the eigenvectors by using laplacian matrix, and use the eigenvector with the second smallest eigenvalues computed using laplacian matrix to bipartition the graph using Eqn. (33) given in [95].

For EM algorithm, at each pixel in an image, we compute a d-dimensional feature vector that encapsulates intensity and texture information. EM algorithm assumes that a segment is chosen with a probability, and models the density associated with that segment as a Gaussian probability distribution function, with parameters (μ, σ), that depend on the chosen segment. This is known as a Gaussian mixture model [96]. The EM tool [34] also yields the cluster mean and covariance, for a user-defined number of clusters and number of iterations. Note that varying the number of clusters and the number of iterations influences the computation time and the quality of results. In our work, we randomly initialize the number of clusters and retain the meaningful number of clusters after couple of iterations.

3.2.5. Image Data Set

The image database includes the two image modalities – gadolinium – enhanced T1, T2 and FLAIR from 10 patients with PF tumors as shown in Table 1. All the images are sampled by 1.5 Tesla Siemens Magnetom scanners. The slice thickness is 5mm, with the slice gap of 1mm, the field-of-view (FOV) of 210x210 mm^2 and the image matrix of 256x256 pixels. The scan parameters for T1- weighted image are: TR=168ms, TE=8ms, flip angle=90 degrees; the scan parameters for T2-weighted image are: Turbo Spin Echo, TR=6430, TE=114ms, 14 echoes per TR.

Table 3.1 MR image data statistics

Patient	Field Strength (Tesla)	Number of tumors	Name of tumor	Number of images with visible tumors	T1 modality			T2 modality		Flair modality	
					Total number of images in a sequence	Tumor visibility	Contrast agent(gadolinium Enhanced)	Total number of images in a sequence	Tumor visibility	Total number of images in a sequence	Tumor visibility
1	1.5	Single	Medulloblastoma	9	35	Good	Applied	35	Good	35	Medium
2	1.5	Single	Medulloblastoma	9	35	Medium	Applied	35	Good	35	Medium
3	1.5	Single	Medulloblastoma	9	35	Medium	Applied	35	Medium	34	Good
4	1.5	Single	Medulloblastoma	8	37	Medium	Applied	36	Medium	36	Medium
5	1.5	Single	Medulloblastoma	9	35	Medium	Applied	35	Medium	34	Good
6	1.5	Single	Astrocytoma	8	40	Good	Applied	40	Good	34	Medium
7	1.5	Single	Astrocytoma	6	27	Medium	Applied	27	Medium	25	Medium
8	1.5	Single	Astrocytoma	8	37	Poor	Applied	38	Medium	35	Medium
9	1.5	Single	Astrocytoma	7	21	Medium	Poor	26	Medium	25	Good
10	1.5	Single	Astrocytoma	9	28	Medium	Applied	27	Medium	26	Good

3.3. Results

3.3.1. Feature Extraction and Selection

We compute intensity, fractal dimension, fractal wavelet and shape features in all MR images for ten patients. We first divide the images into 8 X 8 sub – images and obtain the corresponding features using PTPSA, mBm and level set algorithms respectively. Note in our previous work, we show that the effectiveness of fractal algorithms improve by dividing images into 8x8 sub- images for local detection of tumor

[30]. We then obtain the normalized mean value of the FD, mBm, intensity and shape features for both tumor and non-tumor regions for each MRI slice.

For robust identification of effective features we obtain feature selection using three different techniques such as PCA, boosting and KLD. The PCA [53] offers the feature ranking of the distance of PCA eigenvalues algorithm as shown in Table 3.2. Table 3.2 shows the PCA values for five even numbered patients for example.

Table 3.2 Feature ranking using PCA

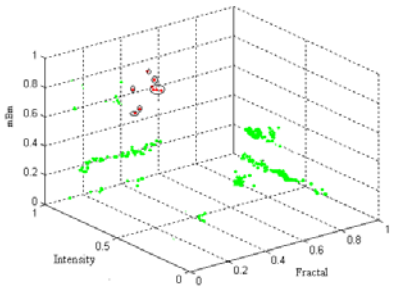
Patient	Features	Distance of eigenvalues for PCA in T1 (feature ranking)	Distance of eigenvalues for PCA in T2 (feature ranking)	Distance of eigenvalues for PCA in FLAIR (feature ranking)	Patient	Features	Distance of eigenvalues for PCA in T1 (feature ranking)	Distance of eigenvalues for PCA in T2 (feature ranking)	Distance of eigenvalues for PCA in FLAIR (feature ranking)
2	Intensity	1.83 (2)	**1.77 (1)**	1.80 (2)	6	Intensity	**1.77 (1)**	**1.71 (1)**	**1.74 (1)**
	mBm	**1.86 (1)**	1.72 (3)	**1.85 (1)**		mBm	1.74 (2)	1.67 (3)	1.72 (2)
	Shape	1.81 (3)	1.74 (2)	1.79 (3)		Shape	1.68 (3)	1.70 (2)	1.68 (3)
	Fractal	1.79 (4)	1.71 (4)	1.77 (4)		Fractal	1.67 (4)	1.66 (4)	1.65 (4)
4	Intensity	1.72 (2)	**1.75 (1)**	1.70 (3)	8	Intensity	1.72 (3)	**1.74 (1)**	1.72 (3)
	mBm	**1.74 (1)**	1.67 (3)	**1.77 (1)**		mBm	**1.80 (1)**	1.70 (3)	**1.76 (1)**
	Shape	1.67 (3)	1.71 (2)	1.75 (2)		Shape	1.77 (2)	1.72 (2)	1.74 (2)
	Fractal	1.66 (4)	1.64 (4)	1.66 (4)		Fractal	1.70 (4)	1.67 (4)	1.65 (4)
10	Intensity	1.71 (2)	**1.73 (1)**	1.74 (2)					
	mBm	**1.75 (1)**	1.67 (3)	**1.76 (1)**					
	Shape	1.69 (3)	1.69 (2)	1.70 (3)					
	Fractal	1.65 (4)	1.65 (4)	1.68 (4)					

Table 3.2 shows that for both T1and FLAIR modalities mBm performs the best for all the patients (except patient 6). For T2 modality, intensity ranks first for all the patients. The second method for feature selection is boosting [22]. The boosting method offers feature ranking in decreasing order using the F scores as shown in Table 3.3 for the same even numbered patients as shown in Table 3.2. Table 3.3 shows that for T1 and FLAIR modality, mBm performs the best. Similarly for T2 modality, intensity ranks first for all the patients.

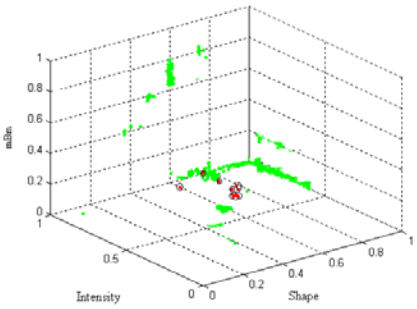
Table 3.3 Feature ranking using F – scores in boosting method

Patient	Features	F-values for T1 (feature ranking)	F-values for T2 (feature ranking)	F- values for FLAIR (feature ranking)	Patient	Features	F values for T1 (feature ranking)	F-values for T2 (feature ranking)	F-values for FLAIR (feature ranking)
2	Intensity	14.4 (2)	**17.7 (1)**	14.8(2)	6	Intensity	18.2 (3)	**18.4 (1)**	18.2 (2)
	mBm	**15.6 (1)**	16.2 (3)	**15.5 (1)**		mBm	**18.7 (1)**	17.6 (3)	**18.4 (1)**
	Shape	13.8 3)	16.8 (2)	13.8 (3)		Shape	18.4 (2)	18.1 (2)	16.8 (3)
	Fractal	13.2 (4)	15.4 (4)	12.3 (4)		Fractal	16.7 (4)	17.3 (4)	16.5 (4)
4	Intensity	18.2 (2)	**18.5 (1)**	16.4 (3)	8	Intensity	18.2 (3)	**17.8 (1)**	17.2 (3)
	mBm	**18.4 (1)**	15.7 (3)	**17.7 (1)**		mBm	**18.0 (1)**	17.4 (3)	**17.6 (1)**
	Shape	16.7 (3)	17.1 (2)	17.5 (2)		Shape	17.7 (2)	17.6 (2)	17.4 (2)
	Fractal	15.6 (4)	14.4 (4)	16.6 (4)		Fractal	16.7 (4)	16.7 (4)	16.5 (4)
10	Intensity	17.3 (2)	**17.6 (1)**	17.4 (2)					
	mBm	**17.7 (1)**	16.7 (3)	**17.6 (1)**					
	Shape	17.2 (3)	17.4 (2)	17.0 (3)					
	Fractal	16.5 (4)	16.5 (4)	16.8 (4)					

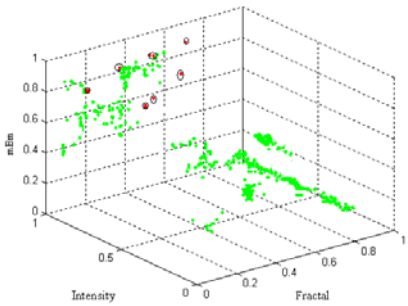
Finally, we obtain KLD plots for all the three MR image modalities per patient. Figure 3.2 shows results in T1, T2 and FLAIR modalities for patient #8 as an example. Figures 3.2 (a) and (c) show that as the entire tumor cluster is located in the mBm plane. Thus, mBm can be used to effectively discriminate between the PF tumors and non-tumor tissues in T1and FLAIR MRI. Figures 3.2(b) shows that intensity is necessary to isolate tumor cluster in T2. This similar trend is noted for all the ten patients in our database. Fig.3.2 clearly provides more effective separation of tumor features. Table 3.4 summarizes our qualitative KLD feature plot observations for all ten patients in all the modalities. We observe that in T1 and FLAIR modality, mBm is the most effective feature for PF tumor segmentation. For T2 modality, intensity is the best features. In order to obtain a more quantitative measure of feature effectiveness, we obtain the entropy (or information gain) for all the four features in T1, T2 and FLAIR modalities respectively. We then rank these entropies in decreasing order.



(a)



(b)



(c)

Fig. 3.2 KLD results showing the separability of features for (a) T1 modality; (b) T2 modality; (c) FLAIR modality for patient#8. Encircled dots show tumors and the rest shows non –tumor.

Table 3.4 Summary of qualitative observation of feature ranking using KLD in T1, T2 and FLAIR MRI.

Patient	Best features using KLD for segmentation inT1 (feature ranking)	Best features using KLD for segmentation in T2 (feature ranking)	Best features using KLD for segmentation in FLAIR (feature ranking)	Patient	Best features using KLD for segmentation inT1 (feature ranking)	Best features using KLD for segmentation in T2 (feature ranking)	Best features using KLD for segmentation in FLAIR (feature ranking)
1	mBm (1) Intensity, fractal (2)	Intensity, shape (1) Intensity (2)	mBm (1) Intensity (2)	6	mBm (1) Intensity, shape (2)	Intensity (1) Intensity, shape (2)	mBm (1) Intensity, shape (2)
2	mBm (1) Intensity, shape (2)	Intensity (1) Intensity, shape (2)	mBm (1)	7	mBm (1)	Intensity (1)	mBm (1)
3	mBm (1) Intensity, shape (2)	Intensity (1)	mBm (1)	8	mBm (1) Intensity, fractal (2)	Intensity (1)	mBm (1)
4	mBm (1)	Intensity (1) Intensity, shape (2)	mBm (1) Intensity, shape (2)	9	mBm (1)	Intensity, shape (1)	mBm (1)
5	mBm (1)	Intensity (1) Intensity, shape (2)	mBm (1)	10	mBm(1)	Intensity (1)	mBm (1)

Table 3.5 Summary of entropy based feature ranking in T1, T2 and FLAIR modalities.

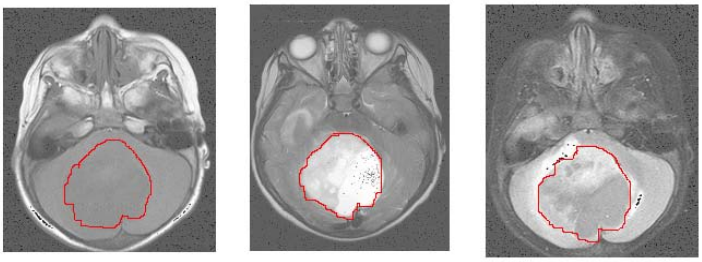
Patient	Features	Entropy values for T1 (feature ranking)	Entropy values for T2 (feature ranking)	Entropy values for FLAIR (feature ranking)	Patient	Features	Entropy values for T1 (feature ranking)	Entropy values for T2 (feature ranking)	Entropy values for FLAIR (feature ranking)
1	Intensity	0.82 (2)	**0.78 (1)**	0.77 (3)	6	Intensity	0.72 (2)	**0.71 (1)**	0.72 (2)
	mBm	**0.84 (1)**	0.69 (4)	**0.82 (1)**		mBm	**0.77 (1)**	0.67 (3)	**0.74 (1)**
	Shape	0.81 (3)	0.75 (2)	0.79 (2)		Shape	0.68 (3)	0.70 (2)	0.68 (3)
	Fractal	0.75 (4)	0.70 (3)	0.75 (4)		Fractal	0.67 (4)	0.66 (4)	0.65 (4)
2	Intensity	0.83 (2)	**0.77 (1)**	0.80 (2)	7	Intensity	0.64 (3)	**0.67 (1)**	0.63 (2)
	mBm	**0.86 (1)**	0.72 (3)	**0.85 (1)**		mBm	**0.68 (1)**	0.63 (3)	**0.70 (1)**
	Shape	0.81 (3)	0.74 (2)	0.79 (3)		Shape	0.66 (2)	0.64 (2)	0.62 (3)
	Fractal	0.79 (4)	0.71 (4)	0.77 (4)		Fractal	0.60 (4)	0.59 (4)	0.60 (4)
3	Intensity	0.78 (2)	**0.78 (1)**	0.76 (2)	8	Intensity	0.72 (3)	**0.74 (1)**	0.72 (3)
	mBm	**0.80 (1)**	0.75 (3)	**0.81 (1)**		mBm	**0.80 (1)**	0.70 (3)	**0.76 (1)**
	Shape	0.76 (3)	0.77 (2)	0.74 (3)		Shape	0.77 (2)	0.72 (2)	0.74 (2)
	Fractal	0.75 (4)	0.74 (4)	0.69 (4)		Fractal	0.70 (4)	0.67 (4)	0.65 (4)
4	Intensity	0.72 (2)	**0.75 (1)**	0.70 (3)	9	Intensity	0.76 (2)	**0.80 (1)**	0.70 (3)
	mBm	**0.74 (1)**	0.67 (3)	**0.77 (1)**		mBm	**0.78 (1)**	0.75 (3)	**0.78 (1)**
	Shape	0.67 (3)	0.71 (2)	0.75 (2)		Shape	0.73 (3)	0.77 (2)	0.75 (2)
	Fractal	0.66 (4)	0.64 (4)	0.66 (4)		Fractal	0.70 (4)	0.72 (4)	0.67 (4)
5	Intensity	0.71 (2)	**0.73 (1)**	0.67 (3)	10	Intensity	0.71 (2)	**0.73 (1)**	0.74 (2)
	mBm	**0.78 (1)**	0.68 (3)	**0.75 (1)**		mBm	**0.73 (1)**	0.67 (3)	**0.76 (1)**
	Shape	0.68 (3)	0.71 (2)	0.69 (2)		Shape	0.69 (3)	0.69 (2)	0.70 (3)
	Fractal	0.66 (4)	0.66 (4)	0.66 (4)		Fractal	0.65 (4)	0.65 (4)	0.68 (4)

Table 3.5 summarizes our ranked entropy results for all ten patients. We observe that in T1 and FLAIR modalities mBm ranks first. In T2 modality, intensity ranks first for all the ten patients. Consequently, using both qualitative KLD features in Table 3.4

and quantitative entropy ranking in Table 3.5, we conclude that mBm is the most effective feature in T1 and FLAIR modalities while intensity is the best for T2 modality. Note these observations are mostly supported by the ranking obtained using boosting method as shown in Table 3.3. For boosting about 70% of feature ranking in T1 modality, 80% of that in FLAIR modality, and 100% of that in T2 modality match between boosting F scores and KLD entropy values. For PCA, about 50% of features ranking in T1 and FLAIR modality and 100% in T2 modality match between PCA and KLD entropy values. However, note that PCA ranking for T1 and FLAIR modalities in Table 3.2 is inconsistent. Furthermore, F scores using boosting in Table 3.3 do not provide consistent feature ranking for T1 and FLAIR modality for all patients (patient # 2,3 and 5 have different F values). Therefore, we use the best features obtained using KLD method such as mBm, intensity and mBm for T1, T2 and FLAIR modalities respectively for subsequent processing.

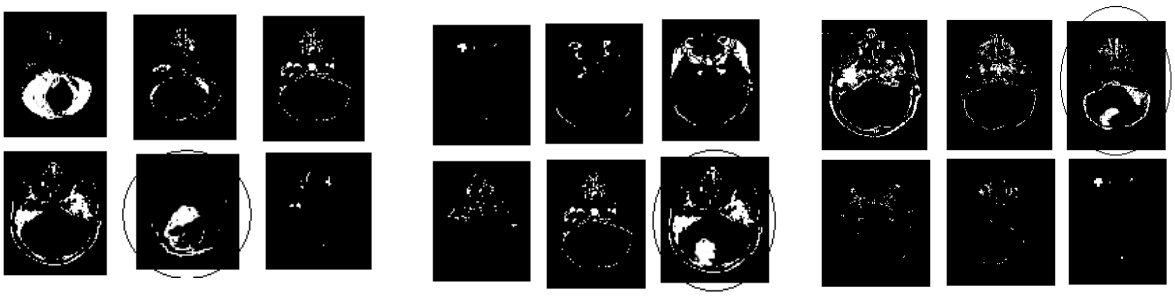
3.3.2. PF Tumor segmentation using selected MRI features

For effective comparison and evaluation, we employ three different tumor segmentation techniques such as top down bottom up, graph cut and EM. Figure 3.3 shows an example for patient #8 in three MRI modalities. The corresponding TDBU segmentation result is shown in Fig. 3.4. Figure 3.4 shows six example clusters each for each MRI modalities. Figure 3.4 shows that the tumor cannot be segmented entirely from the non-tumor regions. We also obtain the summary segmentation results using top down and bottom up method (manual % of area overlap between known ground truth and automated segmentation) for all ten patients as shown in left half of Table 3.6.



(a) (b) (c)

Fig. 3.3 An example MRI slice for (a) T1 modality; (b) T2 modality; (c) FLAIR modality for patient #8. Tumors have been shown using boundary.



(a) (b) (c)

Fig. 3.4 PF tumor segmentation using TDBU method for patient # 8 in (a) T1 modality using mBm; (b) T2 modality using intensity; and (c) FLAIR modality using mBm. Tumors segments are circled.

Next, Figure 3.5 shows the segmentation result for the same patient # 8 using graph cut method. Even though this method offers better segmentation results when compared to those using top down bottom up tumor regions cannot be completely separated from the non-tumor regions. We obtain the summary graph cut segmentation

results (manual % of area overlap between known ground truth and automated segmentation) for all ten patients as shown in right half of Table 3.6.

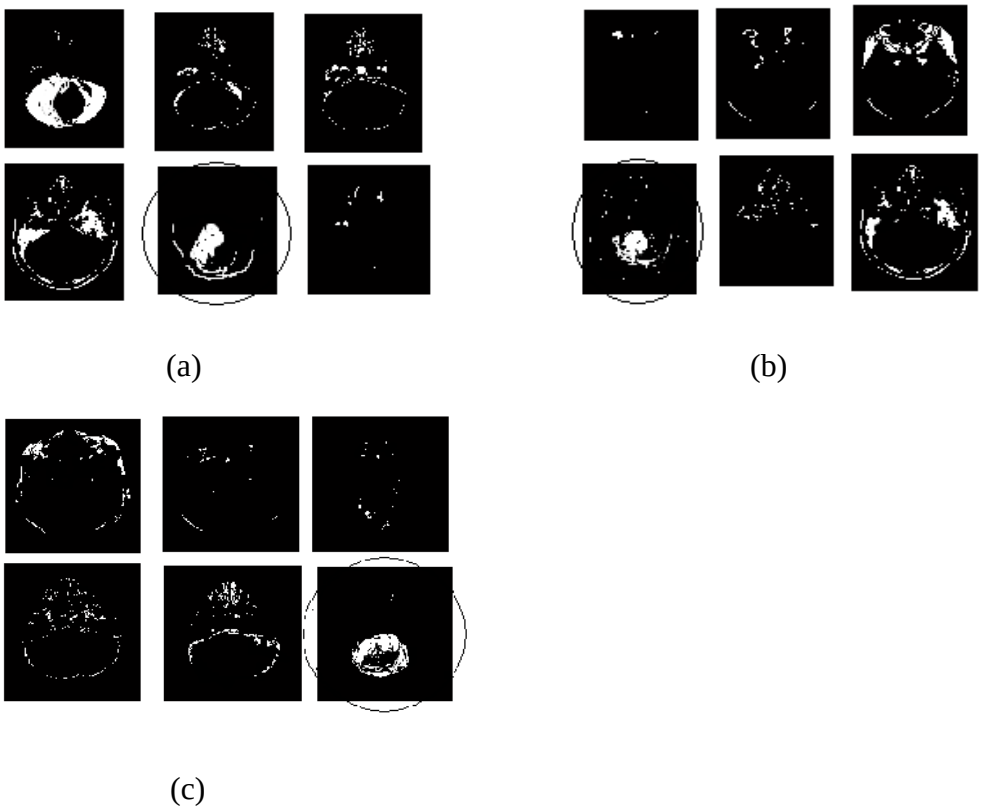


Fig. 3.5 PF tumor segmentation using graph cut for patient # 8 in (a) T1 modality using mBm, (b) T2 modality using intensity, (c) FLAIR modality using mBm respectively. Tumor segments are circled.

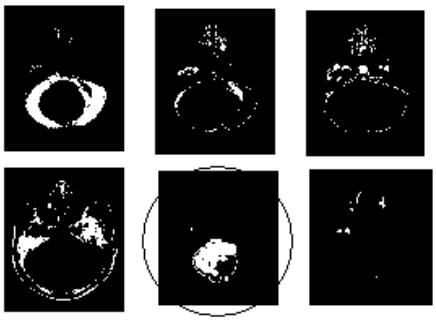
Table 3.6 Summary of tumor segmentation results for top down and bottom up method. The numbers in parenthesis represent the number of images that the tumor region can be clearly segmented vs. the total number of image with visible tumor.

Patient	T1 modality segmentation by TDBU	T2 modality segmentation by TDBU	FLAIR modality segmentation by TDBU	T1 + T2 +FLAIR fusion by TDBU	Patient	T1 modality segmentation by graph cut	T2 modality segmentation by graph cut	FLAIR modality segmentation by graph cut	T1 + T2 +FLAIR fusion by graph cut
	mBm	Intensity	mBm	mBm		mBm	Intensity	mBm	mBm
1	55% (5/9)	66% (6/9)	55% (5/9)	77% (7/9)	1	77% (7/9)	66% (6/9)	66% (6/9)	77% (7/9)
2	77% (7/9)	55% (5/9)	66% (6/9)	77% (7/9)	2	66% (6/9)	66% (6/9)	66% (6/9)	88% (8/9)
3	66% (6/9)	77% (7/9)	55% (5/9)	66% (6/9)	3	55% (5/9)	77% (7/9)	55% (5/9)	77% (7/9)
4	55% (5/9)	77% (7/9)	44% (4/9)	88% (8/9)	4	55% (5/9)	77% (7/9)	55% (5/9)	66% (6/9)
5	77% (7/9)	66% (6/9)	77% (7/9)	77% (7/9)	5	77% (7/9)	88% (8/9)	77% (7/9)	88% (8/9)
6	62% (5/8)	75% (6/8)	62% (5/8)	62% (5/8)	6	62% (5/8)	87% (7/8)	75% (6/8)	75% (5/8)
7	83% (5/6)	66% (4/6)	83% (5/6)	83% (5/6)	7	83% (5/6)	83% (5/6)	83% (5/6)	83% (5/6)
8	75% (6/8)	87% (7/8)	75% (6/8)	87% (7/8)	8	75% (6/8)	87% (7/8)	75% (6/8)	87% (7/8)
9	57% (4/7)	71% (5/7)	71% (5/7)	71% (5/7)	9	71% (5/7)	71% (5/7)	85% (6/7)	85% (6/7)
10	77% (7/9)	77% (7/9)	66% (6/9)	88% (8/9)	10	77% (7/9)	88% (8/9)	66% (6/9)	88% (8/9)
Total	56% (47/83)	72% (60/83)	65% (54/83)	78% (65/83)	Total	69% (58/83)	79% (66/83)	70% (58/83)	81% (67/83)

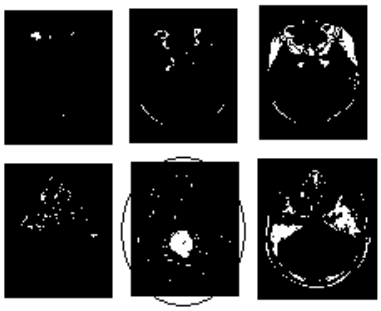
Finally, we obtain tumor segmentation results for the same selected combinations of the features in single modality MR images using EM. Figures 3.6 (a), (b) and (c) show the tumor segmentation using mBm in T1, intensity in T2 and mBm in FLAIR respectively. Comparison among Figs. 3.4, 3.5 and 3.6 shows that EM offers the best tumor segmentation performance. A summary of the complete PF tumor segmentation results using single modality T1, T2 and FLAIR images are shown in the first three columns in Table 3.7. In Table 3.7, we observe that in T1 MRI, mBm offers average segmentation rate (i.e. the number of tumor images segmented vs. total number of images

with visible tumor) of 83%. In T2 modality, intensity yields 84% segmentation rate followed by the combination of intensity and shape with 72%. For FLAIR modality, mBm offers 84% segmentation rate. Comparison of Tables 3.6 and 3.7 confirms that EM is the best segmentation algorithms among the three experimented in this study. Therefore, we employ EM for the next experiment for multiple MRI modality feature fusion.

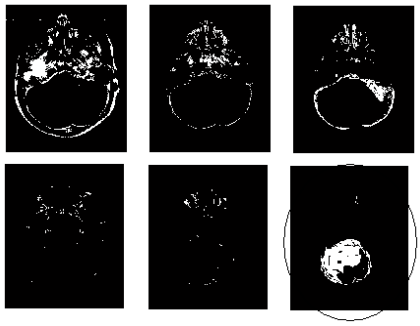
We fuse features in T1, T2 and FLAIR MR modalities for tumor segmentation for each patient. Figures 3.7 (a) show that the entire tumor cluster is located in the mBm plane. Therefore, to find out the best features is mBm; we obtain tumor segmentation using EM in Fig. 3.7(b). In Fig. 3.7 (b), mBm offers better tumor segmentation for fused T1, T2 and FLAIR modalities. We summarize our complete tumor segmentation results for multimodal case in the fourth column in Table 3.7. The results in Table 3.7 suggest that mBm is the best feature for robust PF tumor segmentation for all ten patients in our datasets.



(a)

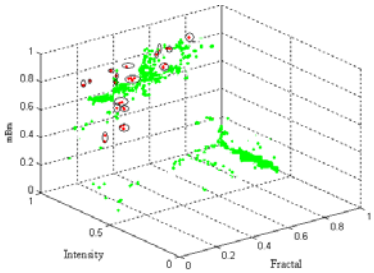


(b)

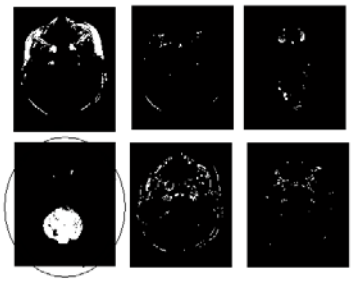


(c)

Fig. 3.6 PF tumor segmentation using EM for patient #8 in (a) T1 image using mBm, (b) T2 image using intensity, (c) FLAIR image using mBm, respectively. Tumor segments are circled.



(a)



(b)

Fig. 3.7 KLD feature fusion of T1, T2 and FLAIR MRI showing separability between (a) intensity, mBm, fractal; (b) Tumor segmentation using EM by fusion of T1, T2 and FLAIR modality for mBm. Tumor segments are circled.

Table 3.7 Summary of tumor segmentation results for EM. The numbers in parenthesis represent the number of images that the tumor region can be clearly segmented vs. the total number of image with visible tumor.

Patient	T1 modality segmentation by EM	T2 modality segmentation by EM	FLAIR modality segmentation by EM	T1 + T2 +FLAIR fusion by EM

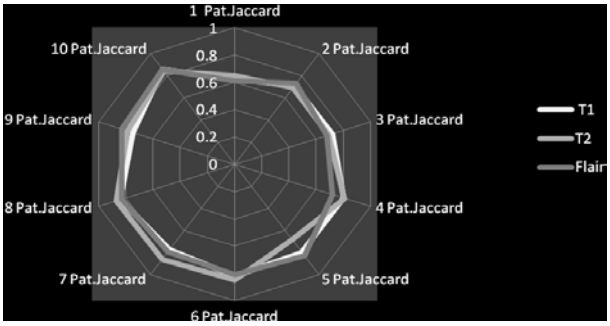
Table 3.7 Summary of tumor segmentation results for EM. The numbers in parenthesis represent the number of images that the tumor region can be clearly segmented vs. the total number of image with visible tumor (cont.)

	mBm	Intensity	Intensity+ shape	mBm	mBm	Intensity +shape
1	77% (7/9)	88% (8/9)	55% (5/9)	88% (8/9)	100% (9/9)	100% (9/9)
2	77% (7/9)	66% (6/9)	66% (6/9)	88% (8/9)	100% (9/9)	77% (8/9)
3	88% (8/9)	77% (7/9)	66% (6/9)	77% (7/9)	100% (9/9)	88% (8/9)
4	88% (8/9)	88% (8/9)	77% (7/9)	77% (7/9)	100% (9/9)	100% (9/9)
5	77% (7/9)	88% (8/9)	77% (7/9)	77% (7/9)	100% (9/9)	88% (8/9)
6	87% (7/8)	87% (7/8)	75% (6/8)	87% (7/8)	100% (8/8)	75% (7/8)
7	83% (5/6)	100% (6/6)	100% (6/6)	83% (5/6)	100% (6/6)	100% (6/6)
8	87% (7/8)	87% (7/8)	75% (6/8)	87% (7/8)	100% (8/8)	87% (7/8)
9	85% (6/7)	85% (6/7)	71% (5/7)	85% (6/7)	100% (7/7)	100% (7/7)
10	77% (7/9)	77% (7/9)	66% (6/9)	88% (8/9)	100% (9/9)	88% (8/9)
Total	83% (69/83)	84% (70/83)	72% (60/83)	84% (70/83)	100% (83/83)	93% (77/83)

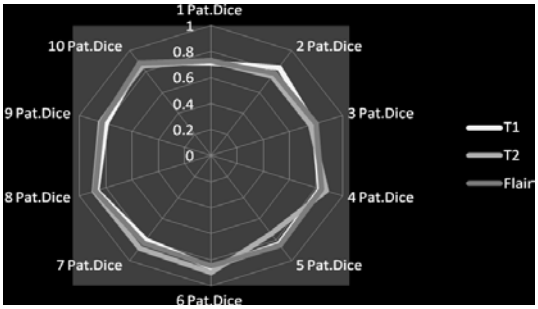
3.3.3. Quality and Robustness of tumor segmentation

In order to verify the quality and robustness of our proposed techniques, we obtain four different similarity measures for automatic computation of overlap between tumors segments obtained using EM and ground truth obtained using manual segmentation by radiologists. Figure 3.8 shows radar plots for four similarity metrics such as Jaccard, Dice, Sneath and Sokal (SS) and Russell and Rao (RR) in T1, T2 and FLAIR modalities for all ten patients, respectively. In each sub plot, for a specific metric the values in y-axis represent overlap coefficient while the axis at each clock location represents patient number. In Figs. 3.8(a) and 3.8(d) both the overall Jaccard and RR overlap is about 60% for all patients. We observe that the Dice overlap in Fig. 3.8 (b) is

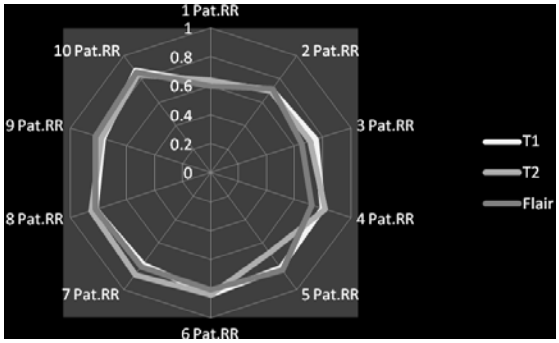
above 80% for all patients. In Figure 3.8(c) SS overlap for nine patients is above 60% except for a dip at 47% for patient # 1 for all modalities. Note these results suggest that our techniques perform better when we compare tumor segments obtained using ground truth to that using our automated segmentation technique as indicated by the Dice metric. However, inclusion of non-tumor segments in the metrics computations, as indicated by both SS and RR metrics, suggests moderate segmentation performance.



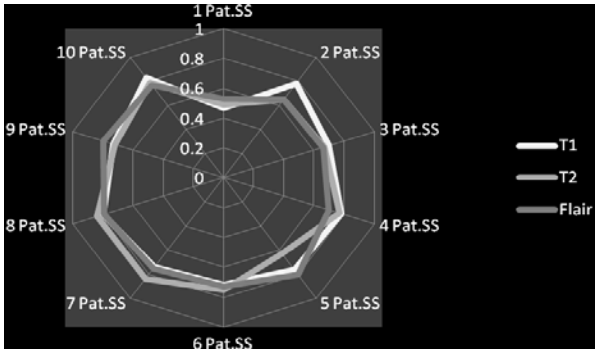
(a)



(b)



(c)



(d)

Fig. 3.8 Plot of similarity metrics for ten patients in 3 modalities for (a) Jaccard; (b) Dice; (c)Sneath &Sokal (SS); and (d)Russell & Rao (RR). The number outside circle shows the patient number from 1 -10.Number 0, 0.2, 0.4, 0.6, 0.8, 1.0 shows the range of metrics.

3.4. Conclusion

We systematically investigate the efficacy of different types of features including texture, (such as FD and mBm) level set shape and intensity for segmentation of PF tumors. For selection of the best feature, we compare four different techniques such as PCA, boosting, KLD and entropy metrics. We implement an integrated mathematical framework for feature selection and ranking using KLD since KLD offers the best feature selection performance for this study. Our KLD feature selection technique shows that mBm is the best feature for both T1 and FLAIR modality while intensity is for T2 modality. In order to obtain robust segmentation of PF tumor in pediatric brain MRI, we compare performance of three different techniques such as bottom up top down, graph cut and EM. We finally select an integrated KLD - EM framework for tumor segmentation since this specific combination offers the best performance among the techniques investigated in this study. The uniqueness of our formal KLD computation takes into account the mean and variance of two different classes expressed in terms of EM. We evaluate robustness of our proposed model using four different similarity metrics and demonstrate the efficacy of our technique using 249 real MR images from ten pediatric patients. Furthermore, we show that fusion of mBm feature in multimodality T1, T2 and FLAIR MRI, can offer 100% PF tumor segmentation for the patient cases studied in this work.

We obtain time estimates of all the steps in this work as shown in Fig. 3.1 such as normalization, feature extraction, features selection and segmentation. In our work, the time taken for normalization is 10 min, extraction of all four features is 30 min, feature selection using KLD is 30 min and segmentation using EM is 40 min respectively for 50

slices/patient on an INTEL[R] Xeon[R] CPU X5355 at 2.66GHz and with 3.00 GB of RAM. Note all these steps can be done offline and made available to aid in a typical clinical setting wherein hundreds of MRI slices may be read by radiologists per day. In the future, we plan to extend our work for automated classification of tumor from non – tumor regions after the PF tumor segmentation. Further, our existing features may not be sufficient to discriminate among the brain tissues such as white matter, gray matter, CSF from tumor and edema. We need to investigate additional features for differentiating among tumor, non-tumor and edema. This will require fundamental work in extending KLD to discriminate multiclass tissues such as brain tissues, tumor, edema and other artifacts in MRI.

4. Information Theoretic Multiclass Feature Selection and Improved Pediatric Brain Tumor Segmentation Robustness Evaluation

4. 1. Introduction

Brain tumor is a leading cause of solid tumor related cancer in children. The tumors are often surrounded by sphere like structure filled with fluid called cyst. Cysts may contain fluid, blood, minerals, or tissue. The cysts are benign growths, but they are sometimes found in parts of the brain that control vital functions. A tumor, on the other hand, consists of a mass of abnormal cells with abnormal growth potential. Cysts have a very thin rim surrounding the fluid for non associated tumors. When tumor has an associated cyst, there is generally a mass, or at least a thickening of the rim, visible on CT or MRI scan [97]. The segmentation of these surrounding tissues such as cyst, necrosis and edema are very difficult due to the surrounding of growth, appearance in MRI, location and size. However a systematic study dealing with their imaging properties such as intensity, shape, selection of best feature and appropriate segmentation technique can be attempted to deal with this problem. To obtain good segmentation we need good features and techniques to select best features from a set. When dealing with tissues such as tumor (T), cyst (C) and non tumor (NT) we need a multiclass selection method which can select the best features for more than two tissues.

Feature selection has been an active area in many different applications. The authors in Refs. [98, 99, 100, 101, 102] discuss various two class feature selection and segmentation techniques in medical imaging domain. Reference [98] presents an information theoretic approach to evaluate the usefulness of each attribute in a feature vector and fuzzy connectedness method for brain tumor segmentation. Principal

component analysis (PCA) and linear discriminant analysis (LDA) are also used for feature selection [99, 100]. A regularization based feature selection to leverage both the sparsity and clustering properties of the features used for uterine cervix image segmentation [101]. Reference [102] describes fuzzy c-mean (FCM) clustering method for segmenting lateral ventricular compartments in brain magnetic resonance imaging (MRI). The method uses Gaussian smoothing to enable fuzzy c-mean (FCM) to create both a more homogeneous clustering result and reduce effect caused by noise.

In our previous works [27,29,30], we studied efficacy of different types of fractal features such as fractal dimension (FD) and multifractional Brownian motion (mBm), as well as intensity and shape factor for brain tumor segmentation. These features are fused using different segmentation techniques such as SOM and EM algorithms to obtain tumor segmentation from the non-tumor tissues. For our prior segmentation results, tumor segments include surrounding tissues such as cyst, edema and necrosis. In order to increase accuracy of tumor segments, there is need to extract cyst and other non tumor tissues from the tumor segments. Different tissue types can be characterized by different features. Therefore, multiclass feature selection is necessary to address multiple tissue segmentation.

The multiclass feature selection is an active research area [103, 104, 105, 106, 107,108, 109]. Reference [103] discusses a prediction risk based feature selection method using multiple classification support vector machines (SVM). The performance of the proposed method is compared with the previous methods of optimal brain damage based feature selection methods using binary SVM. The authors in [104] present multiclass classifier for tissue classification based on gene expression. In order to obtain optimal

gene subset for classification, a genetic algorithm based model-free gene selection method is proposed [105]. A probabilistic neural network technique [106] is compared with that of machine learning methods similar to decision tree and neural network for multi class classification of gene expression data sets. Authors in Ref. [107] describe a multi-class feature selection scheme based on recursive feature elimination for texture classifications. The feature selection scheme is performed in the context of one-against-all least squares support vector machine classifiers. Reference [108] presents a supervised multi-class feature selection approach, which is based on support vector data description. This method suggested utilizes a sequential backward selection algorithm using the accuracy of classifier to decide which feature to be eliminated. A novel layered genetic programming based feature selection is proposed in [109] that use the multiple-population genetic programming. Genetic algorithms have been explored [110] for partitioning the datasets. Further genetic algorithms has been compared with importance score in [111] which is based on greedy algorithm in which the genetic algorithms give a more robust solution at the expense of computational effort.

Many neural network techniques have been used for the selection of multiple features. Bidiwala et al. [50] have proposed neural network for classifying pediatric posterior fossa tumors using clinical and imaging data. The authors in Ref. [112] present a neural network based approach for identifying salient features for classification of diabetic and breast cancer datasets. The augmented error function forces the neural network to keep low derivatives of the transfer functions of neurons when learning a classification task. Cascade Correlation (C2) nets is an internal wrapper feature selection method [113] which selects features and at the same time adds hidden units to the

growing C2 net architecture. A Bayesian neural network [114] with automatic relevance determination priors has been investigate for joint feature selection and classification in computer-aided diagnosis of medical imaging. The authors in [115] have proposed a two-phase filter and wrapper feature selection algorithm to remove redundant or useless features.

The methods described above for multiclass feature selection use neural networks (NN). These NN based feature selection methods are mostly ad hoc. These methods do not offer quantitative measures of features quality. On the other hand, information theoretic approach measures general statistical dependence between variables. Secondly, they are invariant to monotonic transformations performed on the variables, contrary to linear dimension reducers such as principal component analysis. Finally, information theoretic feature selection approach is independent of the decision algorithm, thus reducing computational complexity contrary to genetic algorithms.

In this work, we exploit an information theoretic approach for multiclass feature selection for pediatric brain tumor segmentation. The goal is to select best features for segmentation of tumor (T), cyst (C) and non tumor (NT) tissue classes such that tumor segmentation accuracy can be improved. In our prior work [116], we obtained Kullback – Leibler Divergence (KLD) metric for texture features to discriminate between two classes i.e. tumor and non tumor. We also obtained the entropy metric to cross validate selected features for tumor and non tumor classes. We further showed advantage of KLD when compared to other feature selection techniques such as boosting and PCA for tumor and non tumor tissue segmentation in brain MRI. In this work, we extend the KLD to multiclass feature selection for T, C and NT tissues. The segmentation is obtained by

using a Baye's classifier which offers Baye's error for the different classes for a given feature. We obtain upper bound for the Baye's error and select feature that offers minimum upper bound in a given set. We further obtain the segmentation accuracy by extracting total number of pixels for T, C, NT classes using a Baye's classifier.

4.2. Datasets and Methods

4.2.1 Image Data Set

Our patient database includes the three image modalities such as gadolinium – enhanced T1, T2 and FLAIR from eight patients with pediatric posterior fossa tumors as shown in Table 4.1. All the images are sampled by 1.5 Tesla Siemens Magnetom scanners. The slice thickness is 5mm, with the slice gap of 1mm, the field-of-view (FOV) of 210x210 mm^2 and the image matrix of 256x256 pixels. The scan parameters for T1-weighted image are: TR=168ms, TE=8ms, flip angle=90 degrees; the scan parameters for T2-weighted image are: Turbo Spin Echo, TR=6430, TE=114ms, 14 echoes per TR.

Table 4.1 Datasets for tumor, cyst and non-tumor

Patient	Field Strength (Tesla)	Number of tumors	Name of tumor	Number of images with visible tumors	T1 modality			T2 modality		Flair modality	
					Tumor visibility	Cyst visibility	Contrast agent(gadolinium Enhanced)	Tumor visibility	Cyst visibility	Tumor visibility	Cyst visibility

Table 4.1 Datasets for tumor, cyst and non-tumor (contd.)

1	1.5	Single	Astrocytoma	9	Good	Good	Applied	Good	Good	Medium	Medium
2	1.5	Single	Astrocytoma	9	Medium	Good	Applied	Good	Good	Medium	Medium
3	1.5	Single	Astrocytoma	9	Medium	Medium	Applied	Medium	Medium	Good	Medium
4	1.5	Single	Astrocytoma	8	Medium	Good	Applied	Medium	Good	Medium	Good
5	1.5	Single	Astrocytoma	9	Medium	Good	Applied	Medium	Good	Good	Good
6	1.5	Single	Astrocytoma	8	Good	Medium	Applied	Good	Good	Medium	Good
7	1.5	Single	Medulloblastoma	6	Good	Good	Applied	Medium	Good	Medium	Good
8	1.5	Single	Medulloblastoma	8	Good	Good	Applied	Medium	Good	Medium	Good

4.2.2. Feature selection and segmentation robustness

The overall flow diagram of the method followed is shown in Fig. 4.1. The first step includes the preprocessing stage that minimizes intensity bias in MRI. After intensity normalization we compute the priors for the T, C, and NT tissues. We then extract texture features such as FD using PTPSA algorithm, and mBm using fractal wavelet algorithm in MR images. We use different combinations of these features to Baye's classifier wherein the distance between two classes is computed using KLD, Bhattacharya and JM measures for feature selection. The selected best features are utilized for finding the number of pixels for T, C and NT tissues. These pixels are used as the input to Baye's classifier to obtain the posterior probabilities for respective tissues. We then find segmentation accuracy based on posterior probabilities. We discuss each step below.

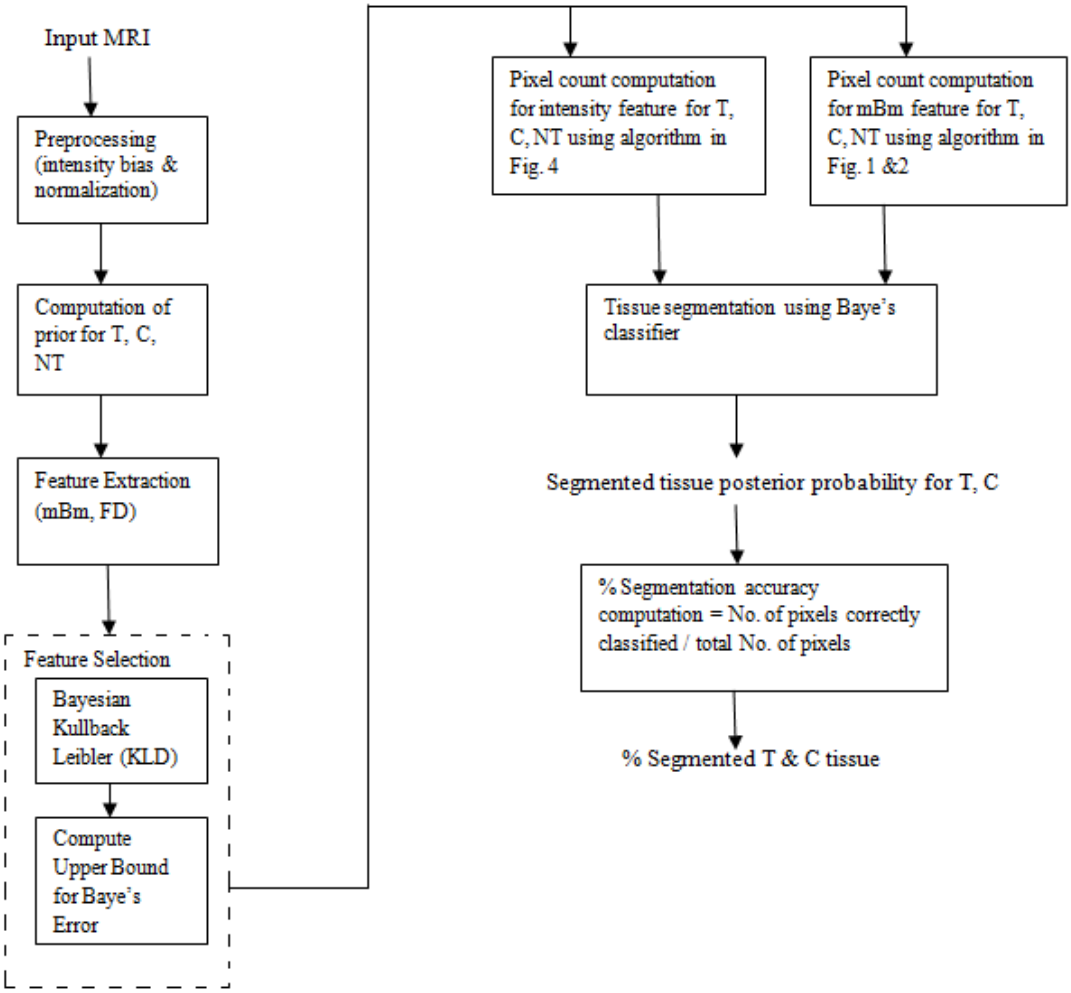


Fig. 4.1 Flow diagram showing feature selection and segmentation accuracy method

4.2.2.1. Image Intensity Normalization

To minimize the intensity bias of the MR image, intensity normalization is used as pre – processing step. In this work, we implement a two- step normalization method [24], wherein the image histograms are modified such that the histograms match a mean histogram obtained using the training data. After applying the normalization method the

intensity values for the same tissue in different MR images fall into a very narrow range in the normalized image.

4.2.2.2. Prior computation

We divide the image into 8X8 sub images and count the number of sub images covered by T, C and NT. We then compute the prior by dividing the number of sub images for T, C and NT to the total number of sub images.

4.2.2.3. Feature Extraction

We extract FD and mBm texture features from the intensity normalized images T1, T2 and Flair modality. We exploit our existing the texture computing algorithms as discussed in [24].

4.2.2.4. Feature selection using Bayesian KLD

We obtain Bayesian KLD metric for all the given features for T, C and NT classes. We then we obtain the upper bound for all the Bayesian KLD metrics using algorithm in Fig.4.2. The upper bound with a lower value corresponds to lower Bayesian KLD metric and is selected as the best feature.

4.2.2.5. Computing accuracy using Baye's classifier

We compute the number of pixels for each T, C, and NT classes respectively. As discussed in Section 3.2 pixel selection for texture feature such as mBm is not easy as intensity feature. Consequently, we obtain pixel count for mBm feature using algorithm

in Figs. 2.2 and 2.3, respectively. Similarly, we obtain pixel count for intensity using algorithm in Fig. 4.3.

4.2.3. Bayesian Kullback Leibler Divergence Criteria for Multiclass feature selection

We extend our prior two class feature selection method in section 3.1 to multiclass using KLD metric. The KLD distance metric is given as,

$$KLD_{ij} = \frac{1}{2}\left\{\sum_{m=1}^{M_\Omega} \alpha_m^{\omega} \log\left(\frac{1}{\sigma_{\omega i}}\right)^2 - 1 + \sum_{m=1}^{M_{\Omega-\omega}} \alpha_m^{\omega,\Omega-\omega}\left[\log(\sigma_{\omega j})^2 + \left(\frac{\sigma_{\omega i}}{\sigma_{\omega j}}\right)^2 + \left(\frac{\mu_{\omega j} - \mu_{\omega i}}{\sigma_{\omega i}}\right)^2\right]\right\} \tag{34}$$

where, ($\mu_{\omega i}$, $\sigma_{\omega i}$) is the mean and variance for first class, ($\mu_{\omega j}$, $\sigma_{\omega j}$) is the mean and variance for the second class and θ is the control parameter. $\alpha^{\omega}{}_{m}$ is a non negative weight, $\sum_{m=1}^{M_\omega} \alpha_m^{\omega} = 1$ and M_ω is the number of features component. Inserting Eqn. (34) in Eqn. (27), we obtain,

$$P_e(\omega_i, \omega_j) = P(\omega_i)\left[1 - Q\left(\frac{\log P(\omega_i)/\log P(\omega_j) - 1/2\left[\log(\sigma_{\omega j})^2 + \left(\frac{\sigma_{\omega i}}{\sigma_{\omega j}}\right)^2 + \left(\frac{\mu_{\omega j} - \mu_{\omega i}}{\sigma_{\omega i}}\right)^2\right]}{\left[\log(\sigma_{\omega j}) + \left(\frac{\sigma_{\omega i}}{\sigma_{\omega j}}\right) + \left(\frac{\mu_{\omega j} - \mu_{\omega i}}{\sigma_{\omega i}}\right)\right]}\right)\right] + P(\omega_j) Q\left(\frac{\log P(\omega_i)/\log P(\omega_j) + 1/2\left[\log(\sigma_{\omega j})^2 + \left(\frac{\sigma_{\omega i}}{\sigma_{\omega j}}\right)^2 + \left(\frac{\mu_{\omega j} - \mu_{\omega i}}{\sigma_{\omega i}}\right)^2\right]}{\left[\log(\sigma_{\omega j}) + \left(\frac{\sigma_{\omega i}}{\sigma_{\omega j}}\right) + \left(\frac{\mu_{\omega j} - \mu_{\omega i}}{\sigma_{\omega i}}\right)\right]}\right) \tag{35}$$

Consider the upper bound for a two class problem for each features based on error probability as discussed in section 3.1. This upper bound is given as,

$$e_{ij} = \left[P(\omega_i) + P(\omega_j)\right] Q\left(\frac{\sqrt{d_{ij}}}{2}\right) \geq P_e\left(\omega_i, \omega_j\right) \tag{36}$$

Note that the optimal decision parameter is computed by fixing the values of KLD_{ij} at the middle point of the KLD distance between the two classes given by KLD_{ij} / 2 in Eqn. (36). The upper bound evaluates two requirements such as the tightness of bound to the error probability and the load for the computation of this bound. The upper bound is provided by the sum of pair - wise errors, computed for all pairs of classes and the sum is given as,

$$E_1 = \sum_{i=1}^{c} \sum_{j>1}^{c} P_e(\omega_i, \omega_j)\,, \tag{37}$$

where $P_e(\omega_i, \omega_j)$ can be computed by Eqn. (27). Considering the pair - wise upper bounds for multiclass problem, the equation can be written as,

$$E_2 = \sum_{i=1}^{c} \sum_{j=1}^{c} e_{ij}\ . \tag{38}$$

The use of E_1 guarantees a better approximation for the error probability, while that of E_2 slightly reduces the computation load [66]. Our algorithm for computing the Baye's upper bound using KLD is given below.

Algorithm for computing upper bound using KLD

1. For each slice k = 1 to N do

   i. *Divide the image into 8x8 sub images.*

   ii. *Extract the textural features – mBm, FD*

   iii. *Compute the value of α using maximum a posterior probability (MAP) in Eqn. (28).*

   iv. *Compute the distance measure d_{ij} for KLD using Eqn. (34).*

   v. *Compute the $P_e(\omega_i, \omega_j)$ using Eqn.(27).*

   vi. *Compute the upper bound using Eqn.(37) or Eqn. (38).*

   vii. *Select the set that has minimum upper bound.*

2. End

Fig.4.2 Algorithm for computing the upper bound using KLD metric

4.2.4. Segmentation accuracy

4.2.4.1. Segmentation Accuracy using Intensity feature

We are interested in obtaining the tumor segmentation accuracy using pixel intensity feature. The segmentation accuracy can be obtained by computing the number of pixels correctly classified using a Baye's Classifier. We first compute the number of pixels for every class such as T, C and NT. We input total numbers of pixels for each class to Baye's classifier and obtain the posterior distributions for each class. We then calculate the number of pixels correctly classified based on posterior value and, hence, the tumor segmentation accuracy. Our algorithm for computing the intensity pixel segmentation accuracy is given below.

Algorithm for computing segmentation accuracy for intensity feature;

1. For each slice k = 1 to N do

   i. *Compute the mean and variance for the whole image for T1, T2 and FLAIR images.*

   ii. *Apply a threshold on basis of the variances obtained for selecting T, C and NT pixels.*

   iii. *Use number of pixels as input to Bayes classifier and obtain posterior probability.*

   iv. *Multiply the total number of pixels by posterior probability to obtain the number of correctly classifies pixels.*

v. *Obtain Segmentation accuracy % (No. of pixels classified / Total no. of pixels) for given class.*

2. End

Fig. 4.3 Algorithm for computing the segmentation accuracy for intensity feature.

4.2.4.2. Segmentation Accuracy using mBm feature

We are also interested in computing the tumor segmentation accuracy using texture features. As discussed in Section 3.1 we cannot directly work with the pixels since fractal texture extraction is a non - linear process. For computing mBm features during extraction process [27] we first find the covariance for each of the sub image given as;

$$\psi_X(s,\tau) = \frac{\sigma_s^2}{2}\left[|s|^{2H(s)} + |s+\tau|^{2H(s)} - |\tau|^{2H(s)}\right], \tag{39}$$

where σ_s^2 is the variance of the mBm process. The expected value of squared – magnitude of the wavelet transform [27] is given as,

$$E\left\{|W_x(s,a)|^2\right\} = \int\int E\{x(s).x(s+\tau)\}\psi_{s,a}(\tau)\psi_{s+\tau,a}(\tau)\,d\tau ds, \tag{40}$$

Substituting the covariance function of the mBm from Eqn. (16) and $\psi_{s,a}(s)$ in Eqn. (15) yields,

$$E\left\{\left|W_x(s,a)\right|^2\right\} = \frac{\sigma_s^2}{2}\int\int\left[\left|s+\tau\right|^{2H(s)}\right]\left|a\right|^{-1}\psi\left(\left|a\right|^{-1}(\tau - s)\right)\psi\left(\left|a\right|^{-1}s\right)d\tau ds \tag{41}$$

In our work, to obtain number of tumor pixels for mBm feature case, we obtain covariance image and decompose the variance image using multiresolution wavelet theory. The resulting decomposed image is divided into sub images of size 8x8. We then compute the wavelet coefficients for all the pixels in the sub images. We obtain the histogram for each sub images given as,

$$r_k = \frac{E\left\{\left|W_x(s,a)\right|^2\right\}}{\sum E\left\{\left|W_x(s,a)\right|^2\right\}} \tag{42}$$

The histogram offers variation in wavelet coefficients for the sub images. The wavelet coefficient values of pixels in the same texture region are similar, and hence we collect these pixels in one group termed as ‘mountain’. On the other hand, the boundary pixels between two texture regions will have a different value termed as ‘valley’. The valleys are computed using the following criteria, if $h(i-j) > h(i)$ and $h(i) < h(i+j)$, with $h(i) = h(i+k), 0 < k < j$, then let $i' = i + (j-1)/2$ and consider pixel ‘i’ as a valley.

We select a suitable threshold for selecting these valleys which in turn yields the boundary for different regions. However, there are some regions which are left out in this process. These regions are merged and resulting regions provide 'seed' for segmentation. We then obtain the interior and exterior pixels using algorithms in Figs. 2.2 and 2.3.

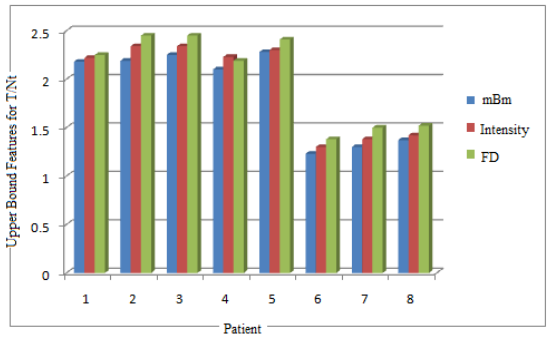
4.3. Results

4.3.1. Multiclass Feature selection using Bayesian KLD

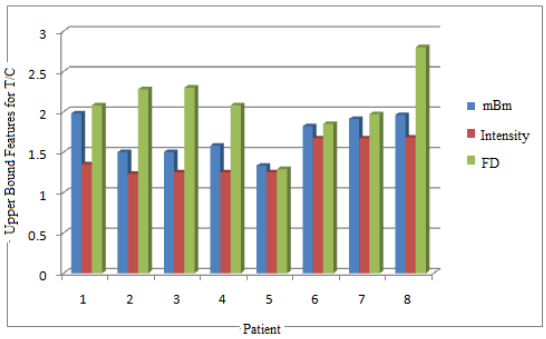
We obtain the Baye's error for the texture and intensity features for T, C and NT classes using Eqn. (11). We then compute the upper bound errors for all three classes using Eqn. (12). Figure 4.4 shows example of upper bound errors for T, C and NT classes using KLD for T1 modality for all eight patients. The upper bound errors for T2 and FLAIR modalities also show similar results.

We observe the minimum value of upper bound from the feature sets for a specific patient. We select the corresponding feature for that upper bound as the best feature for segmentation. In Fig. 4.4 (a), (b) and (c), we observe that for tumor vs. cyst (T/C), cyst vs. non tumor (C/NT) segmentation intensity; and for tumor vs. non tumor (T/NT) segmentation mBm are the best features respectively. Similarly, we obtain best features for T2 and FLAIR modalities. For T2 modality intensity is the best feature for all three classes such as T/C and NT/ C and T/NT. Similarly, for FLAIR modality, intensity is the best feature for T/C and NT/C, while mBm is best for T/NT. Note for comparison we also obtain feature selection using other information theoretic techniques such as Bhattacharya and JM distance measure as shown in Fig.4.5 (a), (b) and (c); and Fig. 4.6 (a), (b) and (c) respectively. Figures 4.4 (a), (b) and (c) shows that KLD is the best metric

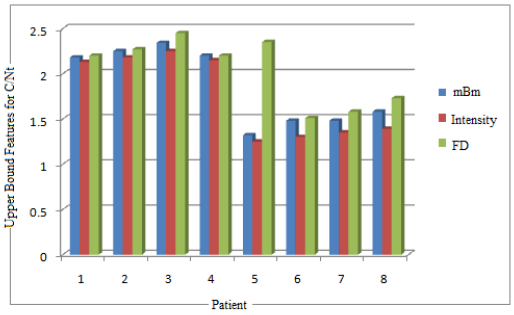
among all metrics for all three T1, T2 and FLAIR MRI modalities. Overall, our feature selection techniques yield intensity and mBm as the best features for T/C, C/NT and T/NT discrimination for all MRI modalities respectively. Therefore, we compute the segmentation accuracy of C, T and NT tissue segmentation using mBm and intensity features for rest of this work.



(a)

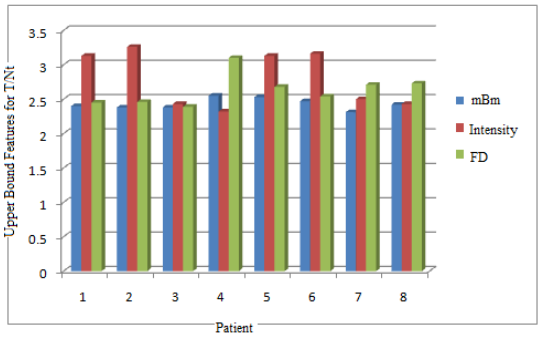


(b)

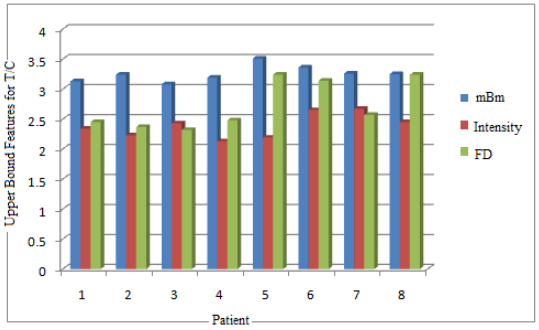


(c)

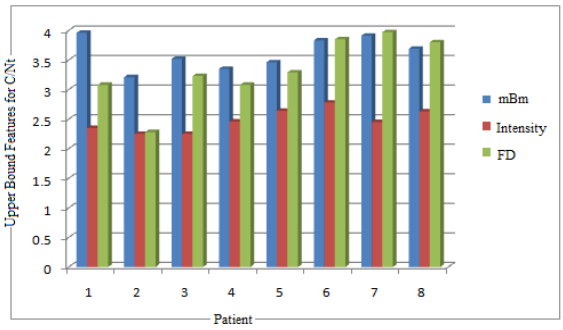
Fig. 4.4 Upper bound for Bayesian KLD framework in T1 modality for (a) T/NT; (b) T/C; and (c) C/NT.



(a)

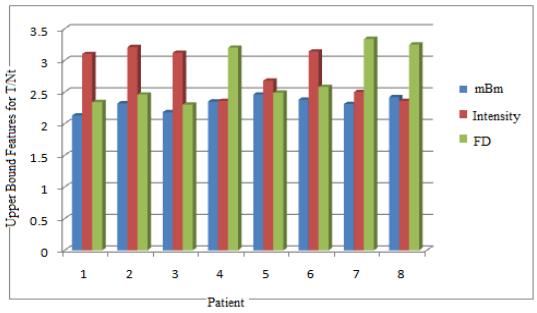


(b)

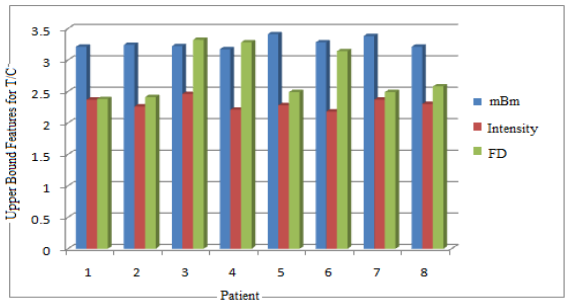


(c)

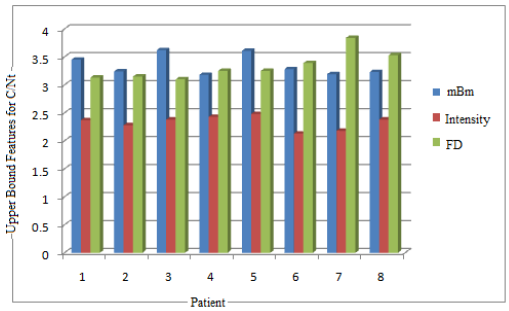
Fig. 4.5 Upper bound for Bhattacharya distance measure in T1 modality for (a) T/NT; (b) T/C; and (c) C/NT.



(a)



(b)



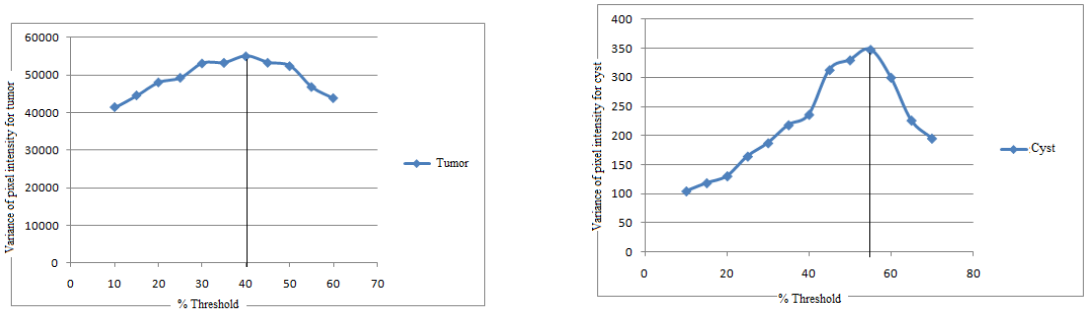
(c)

Fig. 4.6 Upper bound for JM distances measure in T1 modality for (a) T/NT; (b) T/C; and (c) C/NT.

4.3.2. Segmentation Accuracy Computation

4.3.2.1. Segmentation Accuracy for intensity feature

Figures 4.7 (a) and (b) show plots for pixel intensity variance vs. % threshold for an example slice of patient # 2 for tumor and cyst segments respectively. We select the threshold that corresponds to maximum pixel intensity variance in an MRI image. We observe that the appropriate threshold that selects the maximum number of pixles for tumor is 40% while that for cyst is 55%. Note we validate the accuracy of our selected number of pixels for any given tissue type by comparing with that the radiologists segmentation provided for tumor and cyst for each patients. Figure 4.8 shows the segmention of tumor and cyst using intensity as feature for slice # 7 of patient # 2 as an example.



(a) (b)

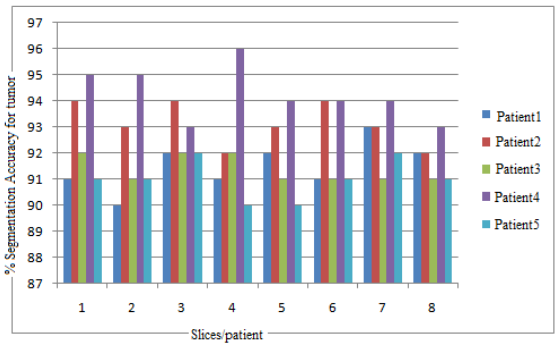
Fig. 4.7 Threshold vs number of pixels selected for (a) tumor; and (b) cyst.

(a) (b) (c)

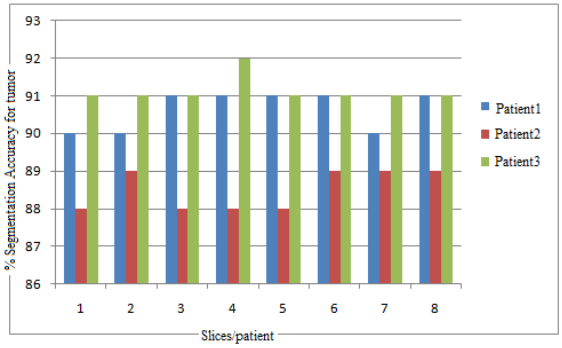
Fig.4.8 (a) MR image for patient #2; (b) Segmented tumor and (b) Segmented cyst for intensity as feature.

We use the selected features obtained in previous step for computing tissue segmentaiton accuracy in a Baye's framework. Note we use five patients to train the network while three patients for testing. Figure 4.9 (a) shows segmentation accuracy vs. slices /patient for T tissue during the training phase of Baye's method in T1 modality.

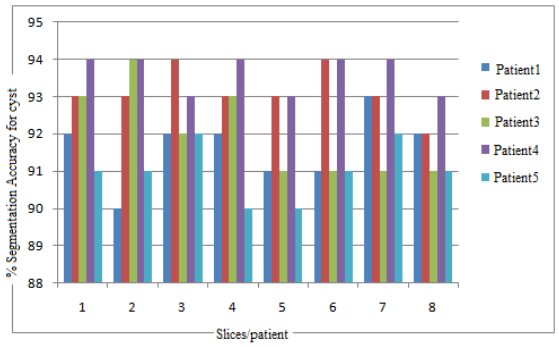
Figure 4.9 (b) shows the corresponding segmentaiton accuracy for tumor tissue during testing phase. Similarly, we obtain the segmentation plots for T1 modality in Figs. 4.9 (c), (d), (e) and (f) for C and NT tissues respectively.We perform the same procedure for T2 and FLAIR modalities for computing segmentaion accuracy for T, C, NT tissues respectively. We observe that the T tissue segmetnation accuracy values for T1 modality are about 95% for training and 90% for testing phases respectively. For C tissue the training and testing values are 94% and 86% respectively as shown in Figs. 4.9 (c) and (d). The training accuracy for NT tissue is 98% while that for testing is 94% as shown in Fig. 4.9 (e) and (f) respectively. Table 4.2 shows the summary for segmnetation accuracy values after training and testing for intensity feature in all the three modalities. Similarly, we obtain training and testing segmentation accuracies for C and NT in T2 and FLAIR moaldities.



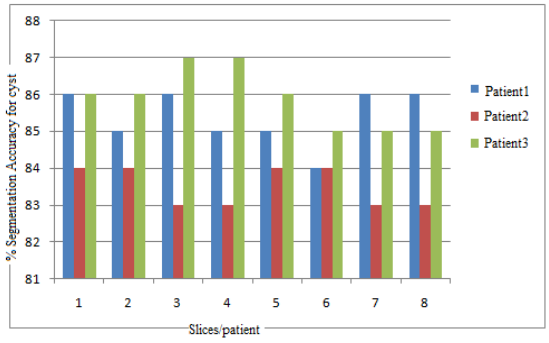
(a)



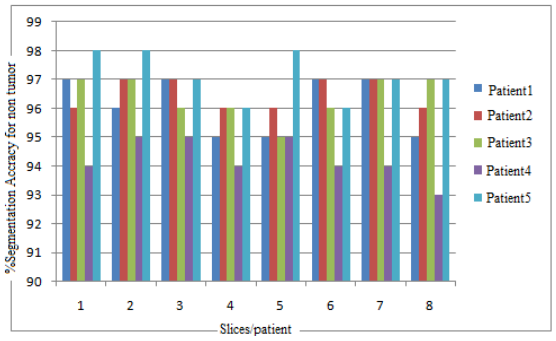
(b)



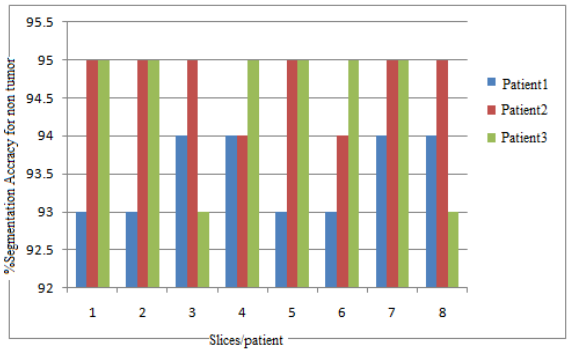
(c)



(d)



(e)



(f)

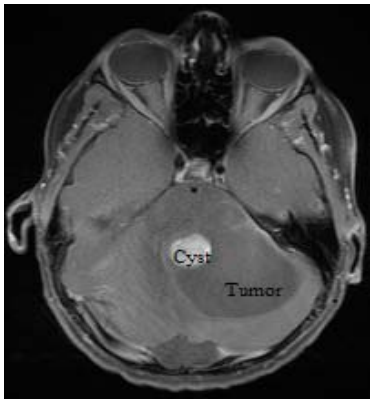
Fig.4.9 Plots of segmentation accuracy vs. slices/patients for (a) Training results for tumor tissue; (b) Testing results for tumor tissue ;(c) Training results for cyst tissue; (d)Testing results for Cyst tissue; (e) Training results for Non tumor tissue; and (f) Testing results for Non tumor tissue in T1 modality.

Table 4.2 Summary of tissue segmentation accuracy using intensity feature for T, C, NT classes

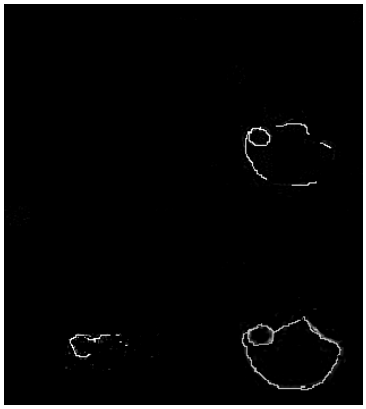
	T1 MRI		T2 MRI		Flair MRI	
Tissue Classes	Training	Testing	Training	Testing	Training	Testing
T	95%	90%	94%	91%	95%	90%
C	94%	86%	94%	94%	94%	88%
NT	98%	94%	97%	97%	96%	92%

4.3.2.2.Segmentation accuracy for mBm feature

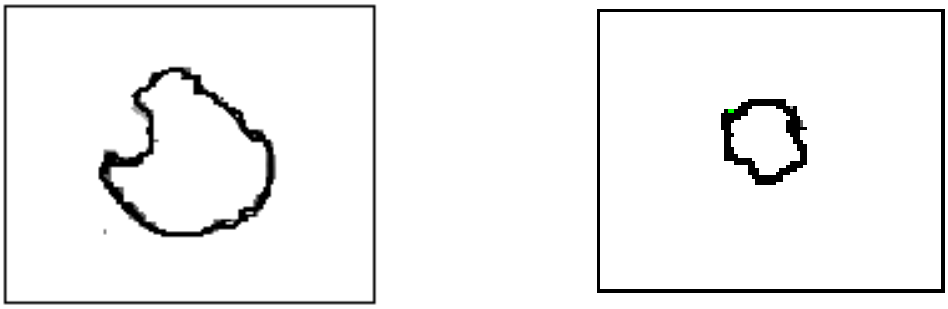
Figure 4.10 (a) shows the MR image of patient # 2 with tumor and the surrounding cyst. Figure 4.10 (b) shows the texture regions after histogram thresholding for subimages. These subimages are obtained after decomposing the image to 1st level. Figure 4.10 (c) and (d) shows the segmentation obtained after defining the boundary using algorithm in Fig.4. 2 for tumor and cyst respectively.



(a)



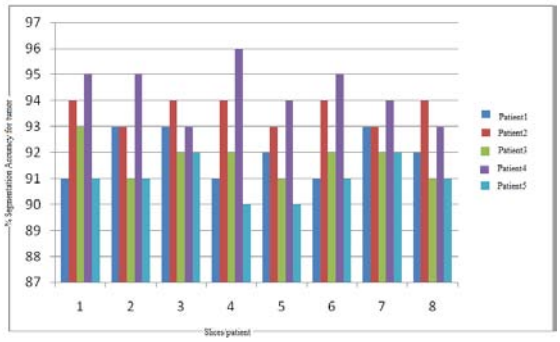
(b)



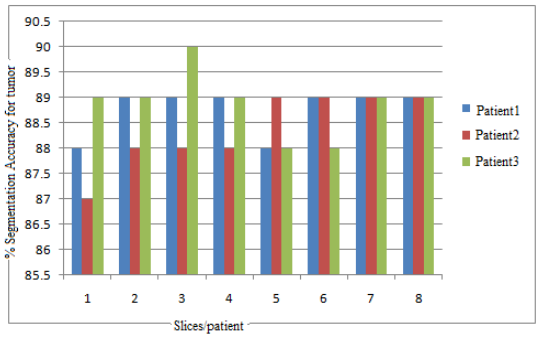
(c) (d)

Fig. 4.10 (a) MR image for patient # 2 in T1 modality. Tumor and cyst are shown by boundary; (b) Tetxure regions obtained after histogram thresholding; (c) Segmented tumor and (d) Segmented cyst after integrating subimages.

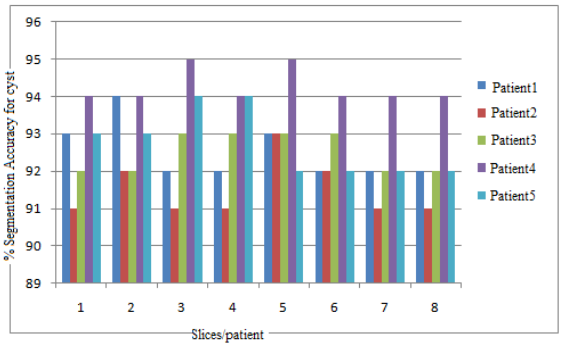
Figure 4.11 (a) shows segmentation accuracy vs. slices /patient for T tissue during the training and testing phase of Baye's method in T1 modality. Figure 4.11(b) shows the corresponding segmentation accuracy for tumor tissue during testing phase. Similarly, we obtain the segmentation plots for T1 modality in Figs. 4.11(c), (d), (e) and (f) for C and NT tissues respectively. We observe that the T tissue segmetnation accuracy value is 93% for training and 89% for testing. For cyst training accuracy for training and testing are 92% and 83% as shown in Figs.4.11(c) and (d). The training and testing accuracy for NT are 94% and 91% shown in Fig. 4.11 (e) and (f). for testing. For NT tissue training accuracies are 93% and 84% for testing. Table 4.3 shows the summary for segmnetation accuracy values after training and testing for intensity feature in all the three modalities. Similarly, we obtain training and testing segmentation accuracies for C and NT in T2 and FLAIR moaldities.



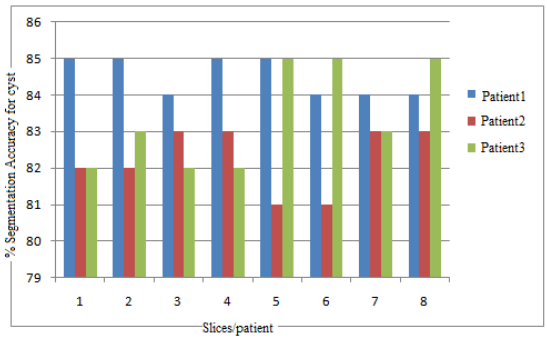
(a)



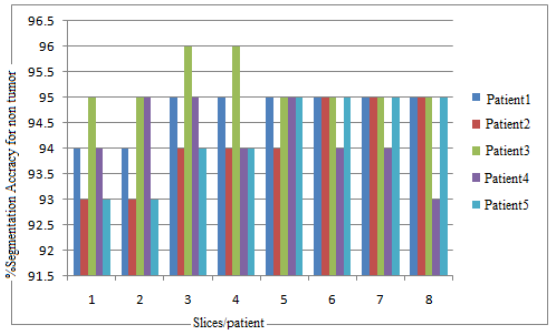
(b)



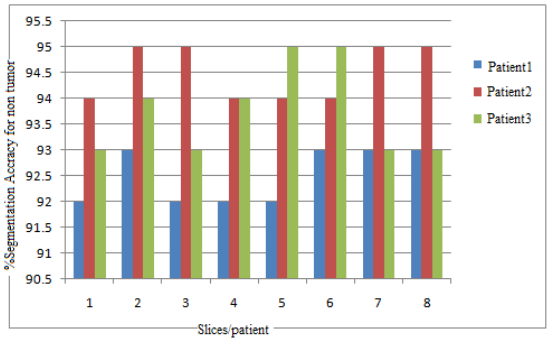
(c)



(d)



(e)



(f)

Fig. 4.11 Plots of segmentation accuracy vs. slices/patients for mBm feature (a) Training results for tumor tissue; (b) Testing results for tumor tissue;(c) Training results for cyst tissue; (d) Testing results for Cyst tissue; (e) Training result for Non tumor tissue; and (f) Testing results for Non tumor tissue in T1 modality.

Table 4.3 Summary of tissue segmentation accuracy using mBm feature for T, C, NT classes

	T1		T2		Flair	
Tissue Classes	Training	Testing	Training	Testing	Training	Testing
T	93%	89%	93%	88%	92%	88%
C	92%	89%	93%	85%	91%	83%
NT	94%	91%	94%	90%	90%	84%

4.3.3 Segmentation robustness

Figure 4.12 shows comparison of tumor segmentation efficacy between our prior two class KLD method [32] and the multiclass KLD method proposed in this work using mBm feature for an example slice for patient 4. Figures 4.12 (a) and (b) show the original image and the segmented tumors and cyst using radiologist's maual segmnetation respectively. Figures 4.12 (c) and (d) show the comparison of T segments respectively. As expected, tumor obtained using the two class KLD method contains cyst. On the other hand, the multiclass KLD method separates the cyst from tumor regions.

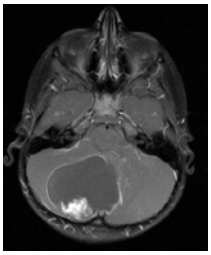
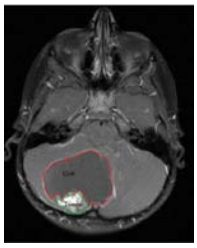
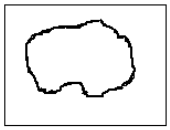
Original Image		Tumor Segmented using 2 class KLD [69]	Tumor segmeted using multiclass KLD proposed in this work	
Raw Image	Manual segmentation using Image J			
 (a)	 (b)	 (c)	Tumor  (d)	Cyst  (e)

Fig. 4.12 Comparison of tumor segmentation results using two class KLD and multiclass KLD

In order to summarize the overall improvement in T segmentation accuracy in this work, we obtain overlap measures. Tables 4.4 and 4.5 show tumor segmentation and coeffcient overlap comparison between our prior work [69] and this study. The similarity coefficients in this study are obtained as the number of pixels correctly segmented for T and C tissues and compare it to the ground truth for T and C region annotated by radiologist. Comparing Tables 4.4 and 4.5 we observe that the tumor segmentation accuracy improves for each patients using our proposed techniques in this work.

Table 4.4 Tumor segmnetation robustness values [31]

Patient	Previous method (Jaccard Coeffcient based on segmented clusters) [31]	Previous method (Dice Coeffcient based on segmented clusters) [31]	Previous method (Sokal & Sneath Coeffcient based on segmented clusters) [31]	Previous method (Russel & Rao Coeffcient based on segmented clusters) [31]
1	80%	83%	81%	84%
2	80%	84%	82%	84%
3	83%	83%	81%	85%
4	82%	83%	83%	84%
5	82%	84%	84%	86%
6	83%	85%	84%	87%
7	83%	85%	85%	87%
8	80%	84%	85%	86%

Table 4.5 Tumor segmentation robustness value using current method

Patient	Current method (Jaccard Coeffcient based on pixels)	Current method (Dice Coeffcient based on pixels)	Current method (Sokal & Sneath Coeffcient based on pixels	Current method (Russel & Rao Coeffcient based on pixels
1	91%	92%	93%	91%
2	94%	94%	93%	94%
3	94%	93%	93%	93%
4	94%	94%	93%	93%
5	93%	94%	93%	95%
6	93%	95%	92%	93%
7	94%	94%	94%	94%
8	93%	95%	94%	94%

4. 4. Conclusion

In this work we investigate an information theoretic multiclass for segmenting T, C and NT tissues respectively. We develop an integrated framework by combining KLD and the upper bound Baye's error for improved pediatirc brain tumor segmentation. We obtain the best features which has minimun upper bound Baye's error from among set of features. Our Baye's KLD approach shows that mBm is the best feature for NT / T segmentation for T1 and FLAIR modalities. In addition, our methods suggests that

intensity is best feature for T / C and C / NT segmentation in T1, T2 and FLAIR modalities. In order to validate the segmentation robustness, we obtain the pixel count in each of the T, C and NT regions using intensity and mBm features. We compute segmentation robustness by comparing the tumor segments to the ground truth provided by radiologists for all eight pediatric patients. In future, we plan to perform more comprehensive work for discriminating other abnormalities such as necrosis and edema for more accurate brain tumor segmnetation.

5. Integrated Framework for Inhomogeneity, Feature Selection and Segmentation in Posterior Fossa Tumors

5.1. Introduction

Segmentation of medical images depends on the structural and intensity characteristics of biological variability. The intensity inhomogeneity [117] [118] can cause a variation in intensity of a particular tissue across the field of view. The most basic tissue-segmentation method is global intensity thresholding. This assumes a voxel intensity can be identified which assigns each voxel into a background class (voxels less intense than the threshold) or a foreground class (voxels more intense than the threshold). Selection of a global threshold may be done in several ways [119] and may not be appropriate in MR images due to intensity inhomogeneity. It may be possible to correct such intensity variation prior to segmentation. An alternative approach is to use local (adaptive) thresholding where the intensity threshold is variable and is computed over sub-images or over a region of interest around each voxel.

Intensity based classification of MR images have proven to be problematic; however, even advanced techniques such as non parametric, multichannel methods have been used. Intra scan intensity inhomogeneities due to RF coils or acquisition sequences e.g. susceptibility artifacts in gradient echo images are a common source of difficulty. In addition, the operating conditions and status of the MR equipment frequently affect the observed intensities causing significant inter scan inhomogeneities which makes it necessary for the manual training on a per scan basis. Reference [120, 121, 122, 123, 124] discusses about some success in correcting intra scan inhomogeneities, such methods require supervision for the individual scans.

Many parametric models have been proposed for solving the inhomogeneities problems in MRI. Ref. [125] discuss about new correction method called PABIC (Parametric Bias field Correction) is based on a simplified model of the imaging process, a parametric model of tissue class statistics, and a polynomial model of the inhomogeneity field. The estimation of the parametric bias field is formulated as a non-linear energy minimization problem using an Evolution Strategy. Li et al. describe about variation level set approach with bias correction for the MR images [126]. In [127], Pham and Prince proposed an energy minimization method for adaptive segmentation and estimation of the bias field. In their method, the smoothness of the bias field is ensured by adding a smoothing constraint term in their objective function, which leads to a highly expensive procedure to solve a space-varying difference equation. Such an expensive smoothing procedure is avoided in some well-known parametric methods by modeling the bias field as a polynomial, which is smooth by nature. However, due to limited approximation capability of polynomials, these methods are not able to approximate bias fields of general profiles, such as those in 7T MR images.

Several works for segmentation due to inhomogeneity has been reported. Kohn et al. [128] observed that inhomogeneity elongates clusters in feature space in the direction of the origin, but that due to the relative positions of the clusters representing brain and cerebra-spinal fluid, the two classes was still separated. Lim et al. [122] proposed a smoothing technique to correct for the inhomogeneity problem: after extraction of the head contour, the intensity values were extended radially towards the image boundaries and smoothed with a Gaussian filter of a large kernel size. They assumed that the resulting blurred image represents one homogeneous region that is only distorted by the

scanner inhomogeneities. The images were corrected with this approximation of the inhomogeneity characteristics. Dawant et al. [123] propose a bias correction method relying on user interaction. A user selects typical sample points of a tissue class as input to the estimation of a parametric bias field. Tincher et al. [129] and Meyer et al. [130] present automatic techniques that fit polynomial functions to pre-segmented regions. The individual fits are combined to find an estimate for a global inhomogeneity field. The procedure relies on a preliminary segmentation into region patches.

So far many integrated approached involving registration and inhomogeneity corrections have been proposed for segmenting in anatomical structures. Warfield et al. [131, 132] combined elastic atlas registration with statistical classification. Elastic registration of a brain atlas helped to mask the brain from surrounding structures. They use “distance from brain boundary” as an additional feature to improve separation of clusters in multi-dimensional feature space. Label fusion methods offer two main advantages: (1) across-subject anatomical variability is better captured than in a single atlas, which can be viewed as a parametric model that typically uses single mode distributions (e.g., Gaussian) to encode anatomical appearance, and (2) multiple registrations improve robustness against occasional registration failures. Authors describe about the incorporation of prior knowledge information into the multiscale framework through a Bayesian formulation [133]. The probabilistic information is based on an atlas prior and on a likelihood function estimated from a manually labeled training set. The significance of new approach is that the constructed pyramid, reflects the prior knowledge formulated. This leads to an accurate and efficient methodology for detection of various anatomical structures simultaneously.

We have not considered registration in this work as registering an MRI with tumor with an atlas will provide erroneous results if coupled with inhomogeneity and feature selection methods. All these methods described either used a parametric model or an energy minimization method for segmentation and inhomogeneity correction. The major drawback of using parametric model is that the inhomogeneity field is estimated only from intensities of one major tissue and then blindly extrapolated over the whole image. The main assumption for gradient based fitting equation is that sufficiently large homogeneous areas are evenly distributed over the entire image so that local gradients of intensity Inhomogeneity can be estimated by local averaging of image intensity gradient. These major drawbacks of these methods are that some adverse image formation may be integrated. These methods are successful only if homogeneous areas are large and distinctive such as the white matter.

This work uses the knowledge of tissue intensity properties and intensity inhomogeneities to correct and segment MR images. We use an EM step in which each iteration utilizes the knowledge of the tissue type to make accurate estimate in next step. In this work, we combine inhomogeneity, feature selection and segmentation in EM framework. In the inhomogeneity step the unknown parameter is bias field B, and the latent variable are the mean and variances for different classes or tissues. For feature selection step the unknown parameter is best feature and the latent variable are the mean and variance for different classes. These mean and variance are iterated and the values of those are allotted in KLD equation for features. In the segmentation round, the missing parameters are the tumor clusters, and the latent variables are the variance of the best features for tumor cluster which is computed in feature selection step. These all steps

come together perfectly in the mathematical model and the algorithms mentioned in the next sections.

5.2. Methods

5.2.1. Mathematical modeling for Inhomogeneity correction, Feature selection and segmentation

In this work, we establish an EM framework for computing the inhomogeneities *B* and feature selection *FS* for MR images *I*. It is difficult to compute these two parameters without considering any hidden variable. We assume segmentation *G* as a hidden variable. When properly defined, the EM framework gives two important guarantees. First, each iteration yields an improved estimate of *(B, FS)* as measured by eqn. (1). Second the algorithm converges to local maxima of the objective function. The conditional probability distribution function describing *I* is given as $P(I|B, FS, G)$. We want to estimate *B* and *FS* from this framework which is given as

$$\left(\hat{B}, F\hat{S}\right) = \arg\max_{B,FS} \log\left(\sum_{G} P\left(B, FS, G|I\right)\right) \tag{43}$$

Next we incorporate the $P(G|I, B', FS')$ where (B', FS') are estimates of $\left(\hat{B}, F\hat{S}\right)$, into eqn. (43) and define $E_{A|B}(f(C)) = P(A|B)f(C)$ to get the following relationship [31]

$$\left(\hat{B}, F\hat{S}\right) = \arg\max_{B,FS} \log\left(\sum_{G} \frac{P\left(B, FS, G|I\right)P\left(G|I, B', FS'\right)}{P\left(G|I, B', FS'\right)}\right)$$

$$= \arg\max_{B,FS} \log E_{G|I,B',FS'}\left(\frac{P\left(B, FS, G|I\right)}{P\left(G|I, B', FS'\right)}\right)$$

The purpose of these operations is to put Eqn. (43) into a form such that we can exploit the bound derived via Jensen's equality. The lower bound of the function is easy to maximize using EM.

$$\log E_{G|I,B',R',FS'}\left(\frac{P\left(B, FS, G|I\right)}{P\left(G|I, B', FS'\right)}\right) \geq E_{G|I,B',R',FS'}\left(\log\frac{P\left(B, FS, G|I\right)}{P\left(G|I, B', FS'\right)}\right) \qquad (44)$$

$$\left(B', FS'\right) = \arg\max_{B,FS} E_{G|I,B',FS'}\left(\log P\left(B, FS, G|I\right) - \log P\left(P|I, B', FS'\right)\right)$$

$$= \arg\max_{B,FS} E_{G|I',B',FS'}(\log P(B, FS, G|I)$$

$$= \arg\max_{B,FS} E_{G|I',B',FS'}\left(\log P\left(B, FS|G, I\right) + \log P\left(G|I\right)\right)$$

$$= \arg\max_{B,FS} E_{G|I',B',FS'}\left(\log P\left(I|G, B, FS\right) + \log P\left(B, FS|G\right) - \log P\left(I|G\right) - \log P(G)\right)$$

$$= \arg\max_{B,FS} E_{G|I',B',FS'}\left(\log P\left(I|G, B, FS\right) + \log P\left(FS|G, B\right) + \log P\left(B|G\right)\right) \qquad (45)$$

Both inhomogeneity and feature selection are affect the segmentation in MRI. But in this work we are assuming them as separate parameters. The optimization procedure decomposes the equation based on the following independence assumptions. First, we assume the independence of *I* with respect to *FS* conditioned on *T* and *B*. We can therefore characterize each anatomical structure with an intensity distribution based on the tissues or classes which is not influenced by the mapping between the atlas and image space. Secondly; we assume *FS* independent of *B* conditioned on *T*. Thirdly; we assume independence of *B* with respect to *T* as the image inhomogeneities are caused by the radio frequency coil of the scanner. Thus, Eqn. (45) simplifies to

$$\left(B', FS'\right) = \arg\max_{B,FS} E_{G|I',B',FS'}\left(\log P\left(I|G,B\right) + \log P\left(FS|G\right) + \log P\left(B\right)\right) \qquad (46)$$

The hidden variable G = {G_1, G_2,……G_n} are the number of segments for each pixel *'x'* denoted by G_x and takes values from the set of k-dimensional unit vectors {e1, e2,…….eK}, where $G_x = eK$, meaning that *'x'* pixels belong to tissue *'k'* or cluster *'k'*. The E step is equivalent to calculating the probability map in the presence of hidden variable *G* and given the estimates of B'_x, FS' for a particular tissue *'k'* using Baye's rule.

$$W_x(k) = \frac{P\left(I_x|G_x(k) = e_k, B'_x, FS'\right)P\left(G_x(k) = e_k|B'_x, FS'\right)}{P\left(I_x|B'_x, FS'\right)} \qquad (47)$$

Adding term $W_x(k)$ to Eqn. (46) simplifies to

$$(B', FS') = \arg\max_{B,FS} \sum_x \sum W_{x(k)} \left(\log P\left(I_x \middle| G_{x=ek}, B\right) + \log P\left(FS \middle| G_{x=ek,}\right) + \log P(B) \right) \quad (48)$$

The M-step maximizes the estimates parameters B' and FS' on probability maps $W_x(k)$.

$$B' = \arg\max_B \sum_x \sum W_{x(k)} \log P\left(I_x \middle| G_{x=ek}, B\right) + \log(B) \quad (49)$$

$$FS' = \arg\max_{FS} \sum_x \sum W_{x(k)} \log P\left(FS \middle| G_{x=ek}\right) + \log(FS) \quad (50)$$

Estimating the intensity inhomogeneities:

In Eqn. (48) the inhomogeneities is defined as

$$P\left(I_x \middle| G_{x=ek}, B\right) = \frac{1}{\sqrt{(2\Pi)^n \left|\gamma_k\right|}} e^{\frac{-1}{2}(I_x - B_x - \mu_k)^G . \gamma_k^{-1} (I_x - B_x - \mu_k)} \tag{51}$$

where γ_k, μ_k are the mean and variances for a particular tissue, I_x is value of intensity feature at pixel x, β_x is the bias field at pixel x for particular tissue or class.

On differentiating Eqn. (48) we obtain

$$0 = \sum_x W_x . \frac{\partial}{\partial B_x} \log P\left(I_x \middle| G_x, B_x\right) + \frac{\frac{\partial P(B)}{\partial B_x}}{P(B)}$$

$$= W_x^G . A_x + \frac{\frac{\partial P(B)}{\partial B_x}}{P(B)}$$

where $A_x = \gamma_k^{-1} . (I_x - \beta_x - \mu_k)$ (52)

turns Eqn .(49) into a closed from solution. In practice we achieve good result by estimating *B* by a low pass filter applied to a weighed residual that depends on *W*, $\left(\mu_{k,} \gamma_k\right)$ and image *I*.

Estimating the feature selection:

In Eqn. (48) the feature selection is defined by using KLD which is explained in Aim 1. Aim 1 also describes about derivation of KLD for two classes.

$$P\left(G_{x=ek}\middle|FS\right)=\frac{1}{2}\left\{\log\left(\frac{1}{\sigma_k}\right)^G-1+\left[\log(\sigma_k)^G+\left(\frac{\sigma_k}{\sigma_m}\right)^G+\left(\frac{\mu_m-\mu_k}{\sigma_k}\right)^G\right]\right\} \tag{53}$$

On differentiating Eqn. (49) we obtain

$$0=\sum_x W_x.\frac{\partial}{\partial FS_x}\log P(G_x|FS)+\frac{\frac{\partial P(FS)}{\partial FS_x}}{P(FS)}$$

$$=W_x^T.C_x+\frac{\frac{\partial P(FS)}{\partial FS_x}}{P(FS)}$$

$$\text{where } C_x=\left(\frac{\sigma_k}{\sigma_m}\right)^G+\left(\frac{\mu_m-\mu_k}{\sigma_k}\right)^G \tag{54}$$

where σ_m, σ_k , μ_m , μ_k are the mean and variance for different tissues or classes.

The segmentation *G* depends on the best feature selected using KLD. The KLD represents the conditional probability for two classes or tissues which are T/NT, T/C and C/NT. The KLD considers the means and variance for the two classes or tissues for a

particular texture feature and these means and variances are updated during M step. The segmentation for different tissues is related with the updating of probability maps which are updated for inhomogeneity and feature selection.

EM applied to Inhomogeneity, Feature selection and Segmentation:

Substituting eqn. (51) and (53) in eq. (47) gives,

$$Wx(k) = \frac{P\left(I_x \middle| G_{x=ek}, B'_x\right) P\left(G_{x=ek} \middle| FS'\right)}{\sum_k P\left(I_x \middle| G_{x=ek}, B'_x\right) P\left(G_{x=ek} \middle| FS'\right)}$$

$$= \frac{\dfrac{\left(\left|\gamma_k\right|\right)^{-0.5} e^{\frac{-1}{2}(I_x - B_x - \mu_k)^G \gamma^{-1}(I_x - B_x - \mu_k)}}{\left(\left|\gamma_m\right|^{-0.5} e^{\frac{-1}{2}(I_x - B_x - \mu_m)^G \gamma^{-1}(I_x - B_x - \mu_m}\right)}}{\dfrac{\sum_k \left(\left|\gamma'_k\right|\right)^{-0.5} e^{\frac{-1}{2}(I_x - B_x - \mu'_k)^G \gamma^{-1}(I_x - B_x - \mu'_k)}}{\sum_m \left(\left|\gamma'_m\right|^{-0.5} e^{\frac{-1}{2}(I_x - B_x - \mu'_m)^G \gamma^{-1}(I_x - B_x - \mu'_m}\right)}} \frac{\dfrac{1}{2}\left\{\log\left(\dfrac{1}{\sigma_k}\right)^G - 1 + \left[\log(\sigma_k)^G + \left(\dfrac{\sigma_k}{\sigma_m}\right)^G + \left(\dfrac{\mu_m - \mu_k}{\sigma_k}\right)^G\right]\right\}}{\dfrac{1}{2}\left\{\log\left(\dfrac{1}{\sigma'_k}\right)^G - 1 + \left[\log(\sigma'_k)^G + \left(\dfrac{\sigma'_k}{\sigma'_m}\right)^G + \left(\dfrac{\mu'_m - \mu'_k}{\sigma'_k}\right)^G\right]\right\}}$$

(55)

The algorithm for computing inhomogeneity, features selection in EM framework is discussed in Fig. 5.1.

Algorithm for computing Inhomogeneity, features selection in EM framework

For

Input MRI scan = I, Pixel x = 1 to N, Tissue class 1 = k for T, Tissue class 2 = m for NT, interval = 0: 60;

1) *Initialize the weight* B' *and* FS' *using Eqn. (51) and (53)*
2) *Iterate E step using Eqn. (55)*
3) *Update the M step as*

$$B = \sum_{k} W_{x(k)} \gamma_k^{-1} (I_x - B_x - \mu_k)\,,\; B = \sum_{m} W_{x(m)} \gamma_m^{-1} (I_x - B_x - \mu_m)$$

$$FS = \sum_{k} W_{x(k)} . \left(\frac{\sigma_k}{\sigma_m} \right)^G + \left(\frac{\mu_m - \mu_k}{\sigma_k} \right)^G$$

4) *Stop at convergence*
5) *Label map* $T_x = \arg\max Wx(k)$.

Fig. 5.1 shows the flow diagram for integrated feature selection and Segmentation.

5.2.2. Feature selection and segmentation in EM framework

The first step includes the preprocessing stage that minimizes the inhomogeneity of the MRI. After inhomogeneity correction we then extract texture features such as FD using PTPSA algorithm, mBm using fractal wavelet algorithm in MR images. We have selected the best features for T vs. NT, T vs. C and C vs. NT using multiclass feature

selection in AIM #2. We feed the best features for the different classes or tissues and obtain subsequent segmentations for the tissues using EM. A detailed explanation for all methods is given below.

5.2.2.1. Estimating Inhomogeneity correction

We use a Bayesian approach to estimate the bias field in MR intensity image [20]. The method assumes a Gaussian distribution for the different tissues or classes.

5.2.2.2. Feature Extraction

After intensity normalization we extract textural features from the normalized images in T1, T2 and Flair modality - FD, mBm.

5.2.2.3. Feature selection using KLD and segmentation

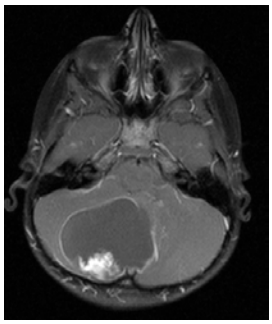
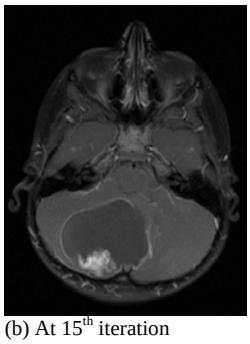
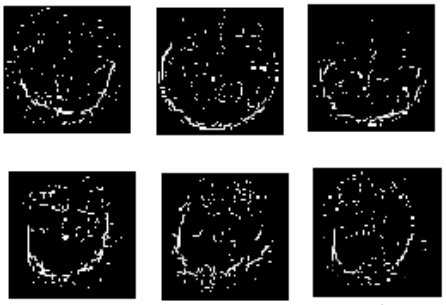
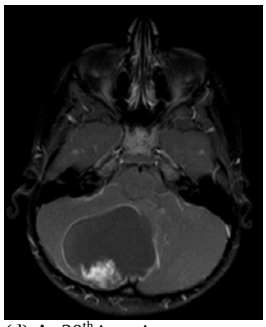
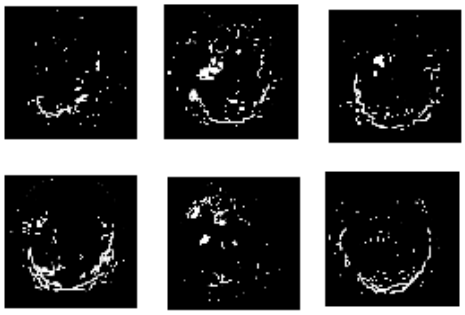
We construct as support map or probability map. These support maps contains the mean and variance associated with a pixel for the best features. We initialize the support map by the estimates of B and FS. We maximize the estimates of B and FS for two classes. The labeling of map for each cluster gives the segmentation for the two classes which in turn has been obtained by the best features.

5.3. Results

Figure 5.2 (a) shows an MR image of patient having tumor in T1 modality. Figure 5.2 (b), (d), (f) and (h) shows the inhomogeneity results at 15th, 30th, 45th and 60th iteration. Fig. 5.2 (c), (e), (g) and (i) show the segmentation results for intensity feature

using integrated KLD- EM algorithm. We observe that good tumor segmentation is obtained at 60th iteration of EM algorithm in cluster no. 5.

Similarly, Fig. 5.3 (b), (d), (f) and (h) shows the inhomogeneity results at 15th, 30th, 45th and 60th iteration. Fig. 5.3 (c), (e), (g) and (i) show the segmentation results for mBm feature using integrated KLD- EM algorithm. We observe that good tumor segmentation is obtained at 60th iteration of EM algorithm in cluster no. 5.

MR Image	Inhomogeneity	Feature selection & Segmentation
 (a)	 (b) At 15th iteration	 (c) Segmentation using intensity feature at 15th iteration
	 (d) At 30th iteration	 (e) Segmentation using intensity feature at 30th iteration

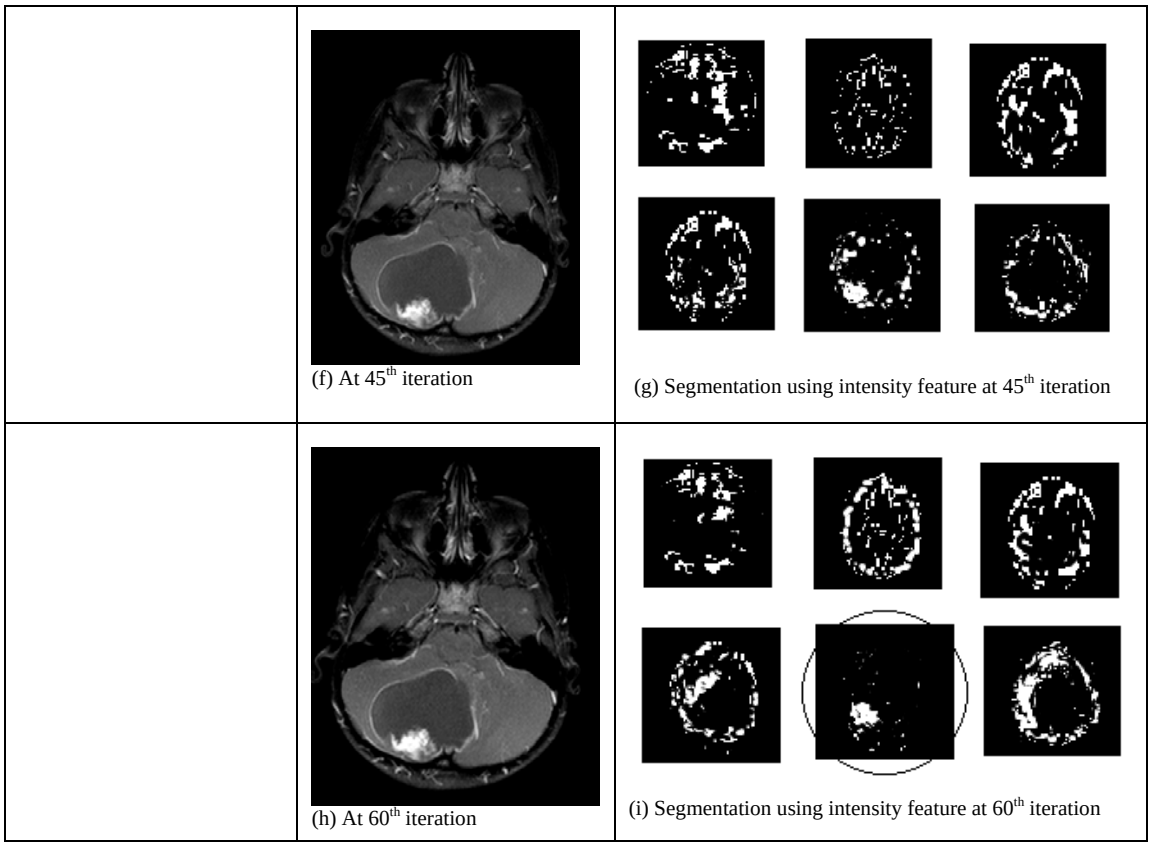
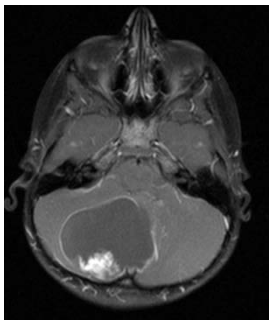
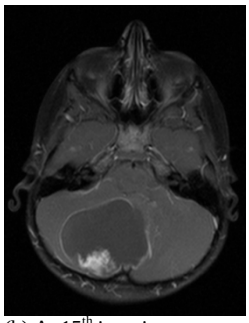
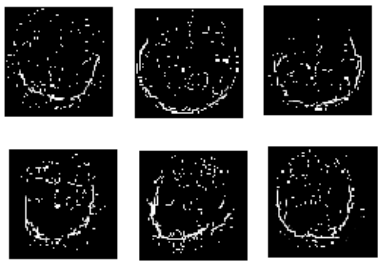
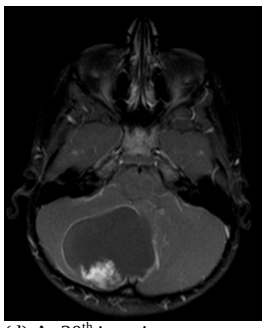
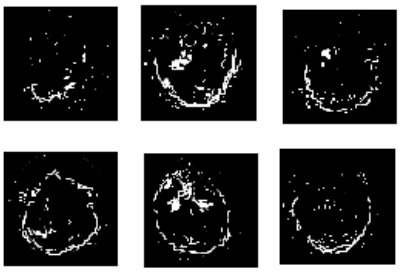
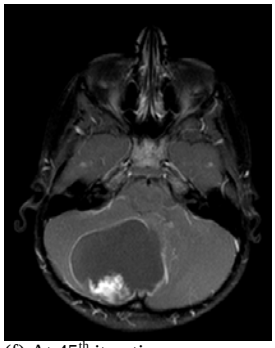
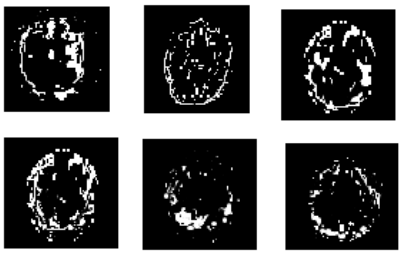


Fig. 5.2 (a) T1 modality for patient #2; (b) inhomogeneity iteration at 15th iteration; (c) segmentation at 15th iteration; (d) inhomogeneity iteration at 30th iteration; (e) segmentation at 30th iteration; (f) inhomogeneity iteration at 45th iteration; (g) segmentation at 45th iteration; (h) inhomogeneity iteration at 60th iteration; (i) segmentation at 60th iteration using KLD –EM framework for T vs. NT for intensity feature in T1 modality.

Figure 5.4 (b), (d), (f) and (h) shows the inhomogeneity results at 15th, 30th, 45th and 60th iteration. Fig. 5.4 (c), (e), (g) and (i) show the segmentation results for FD feature using integrated KLD- EM algorithm. We observe that good tumor segmentation is obtained at 60th iteration of EM algorithm in cluster no. 5.

MR Image	Inhomogeneity correction	Feature selection & Segmentation
 (a)	 (b) At 15th iteration	 (c) Segmentation using mBm feature at 15th iteration
	 (d) At 30th iteration	 (e) Segmentation using mBm feature at 30th iteration
	 (f) At 45th iteration	 (g) Segmentation using mBm feature at 45th iteration

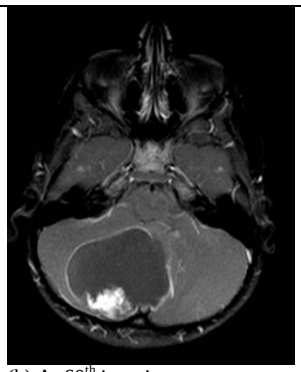
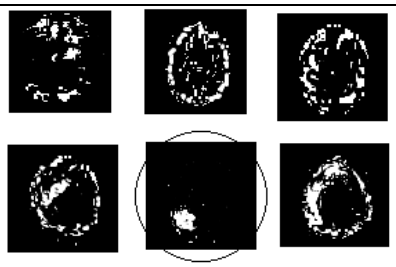
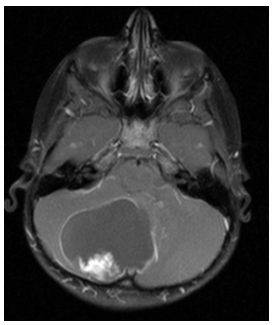
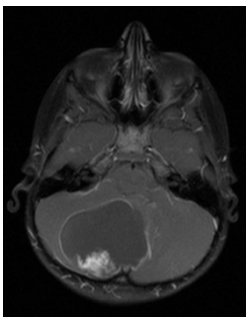
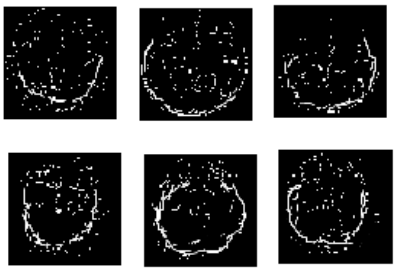
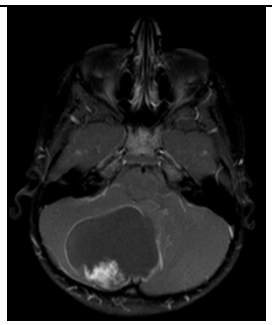
	 (h) At 60th iteration	 (i) Segmentation using mBm feature at 60th iteration

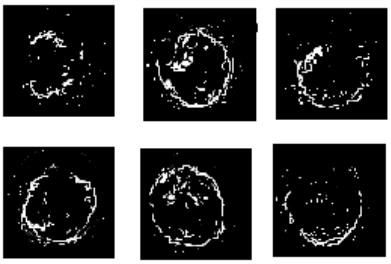
Fig. 5.3 (a) T1 modality for patient #2; (b) inhomogeneity iteration at 15th iteration; (c) segmentation at 15th iteration; (d) inhomogeneity iteration at 30th iteration; (e) segmentation at 30th iteration; (f) inhomogeneity iteration at 45th iteration; (g) segmentation at 45th iteration; (h) inhomogeneity iteration at 60th iteration; (i) segmentation at 60th iteration using KLD –EM framework for T vs. NT for mBm feature in T1 modality.

Note that for features such as mBm and FD we first perform the inhomogeneity correction and then extract texture features offline. After this we perform the automated steps for features selection and segmentation.

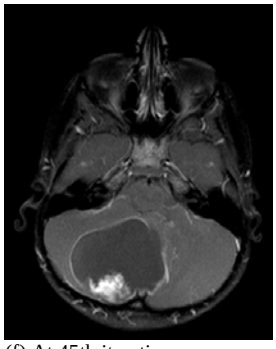
MR Image	Inhomogeneity correction	Feature selection & Segmentation
 (a)	 (b) At 15th iteration	 (c) Segmentation using FD feature at 15th iteration



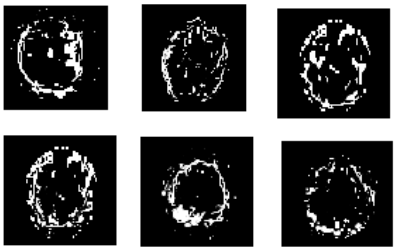
(d) At 30th iteration



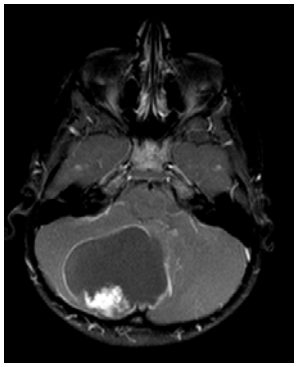
(e) Segmentation using FD feature at 30th iteration



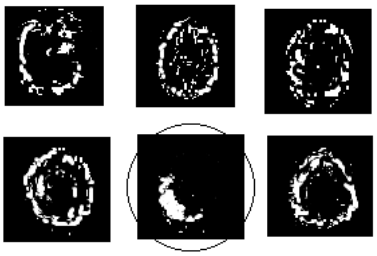
(f) At 45th iteration



(g) Segmentation using FD feature at 45th iteration



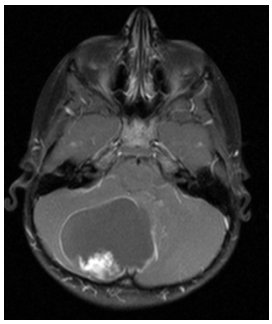
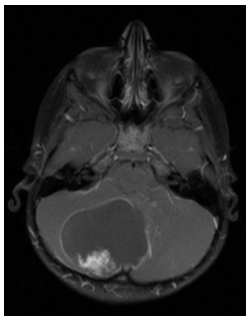
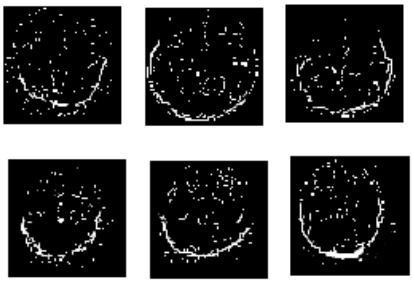
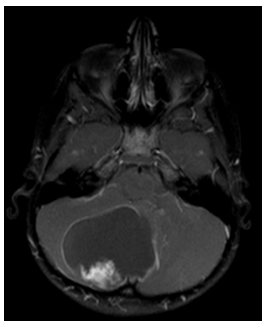
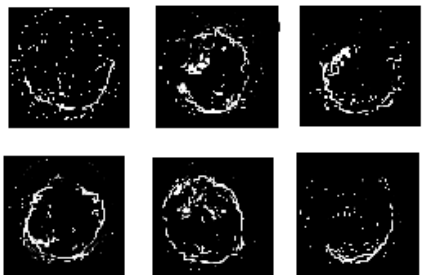
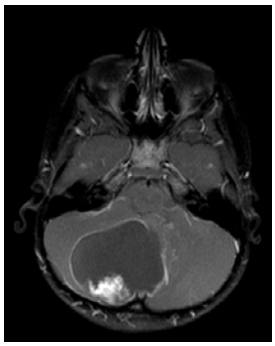
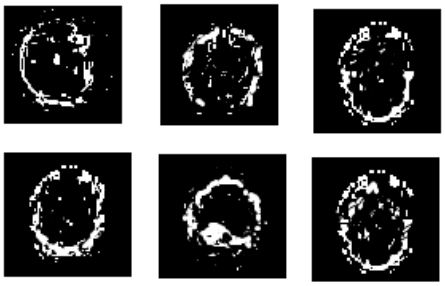
(h) At 60th iteration



(i) Segmentation using FD feature at 60th iteration

Fig. 5.4 (a) T1 modality for patient #2; (b) inhomogeneity iteration at 15th iteration; (c) segmentation at 15th iteration; (d) inhomogeneity iteration at 30th iteration; (e) segmentation at 30th iteration; (f) inhomogeneity iteration at 45th iteration; (g) segmentation at 45th iteration; (h) inhomogeneity iteration at 60th iteration; (i) segmentation at 60th iteration using KLD –EM framework for T vs. NT for FD feature in T1 modality.

Similarly Fig. 5.5 and 5.6 shows segmentation for C from T an NT for intensity features.

MR Image	Inhomogeneity correction	Feature selection & Segmentation
 (a)	 (b) At 15th iteration	 (c) Segmentation using intensity feature at 15th iteration
	 (d) At 30th iteration	 (e) Segmentation using intensity feature at 30th iteration
	 (f) At 45th iteration	 (g) Segmentation using intensity feature at 45th iteration

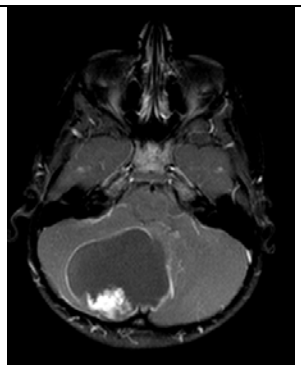
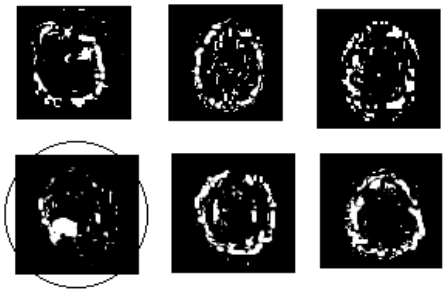
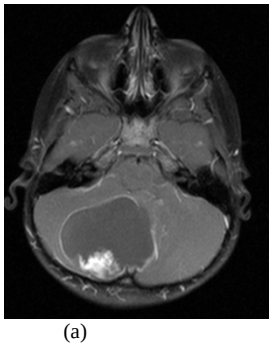
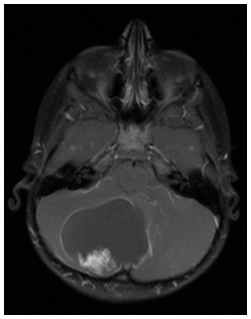
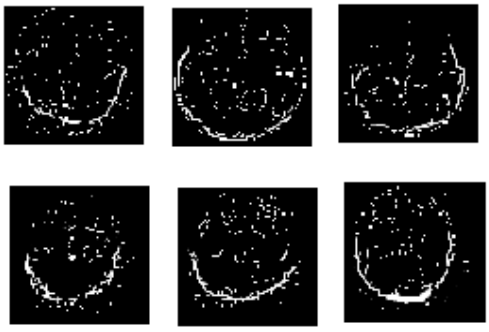
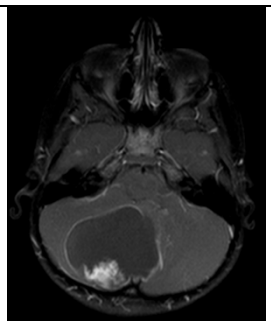
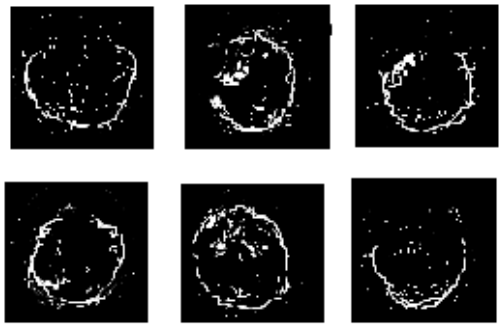
	 (h) At 60th iteration	 (i) Segmentation using intensity feature at 60th iteration

Fig. 5.5 (a) T1 modality for patient #2; (b) inhomogeneity iteration at 15th iteration; (c) segmentation at 15th iteration; (d) inhomogeneity iteration at 30th iteration; (e) segmentation at 30th iteration; (f) inhomogeneity iteration at 45th iteration; (g) segmentation at 45th iteration; (h) inhomogeneity iteration at 60th iteration; (i) segmentation at 60th iteration using KLD –EM framework for C vs. NT for intensity feature in T1 modality.

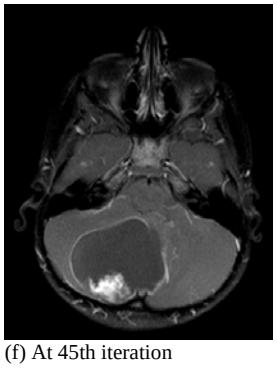
MR Image	Inhomogeneity correction	Feature selection & Segmentation
 (a)	 (b) At 15th iteration	 (c) Segmentation using intensity feature at 15th iteration



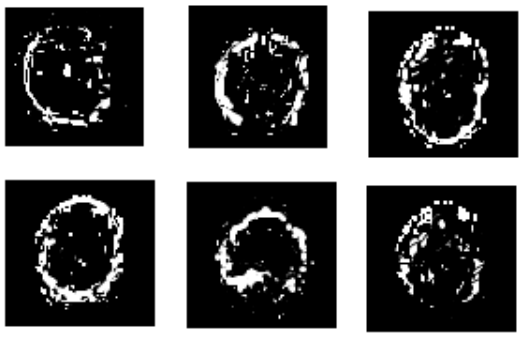
(d) At 30th iteration



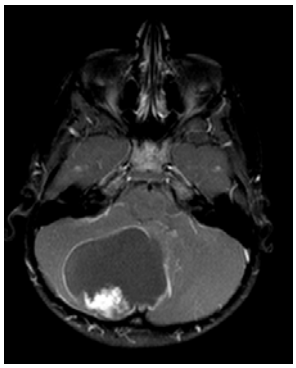
(e) Segmentation using intensity feature at 30th iteration



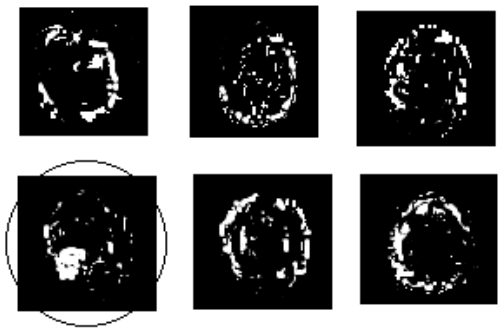
(f) At 45th iteration



(g) Segmentation using intensity feature at 45th iteration



(h) At 60th iteration



(i) Segmentation using intensity feature at 60th iteration

Fig.5.6 (a) T1 modality for patient #2; (b) inhomogeneity iteration at 15th iteration; (c) segmentation at 15th iteration; (d) inhomogeneity iteration at 30th iteration; (e) segmentation at 30th iteration; (f) inhomogeneity iteration at 45th iteration; (g) segmentation at 45th iteration; (h) inhomogeneity iteration at 60th iteration; (i) segmentation at 60th iteration using KLD –EM framework for C vs. T for intensity feature in T1 modality.

Table 5.1 shows the similarity coefficients obtained from the overlap of the segmented tumors to that for the original tumor (ground truth) when KLD and EM have been used separately (AIM#1). Table 5.2 shows the similarity coefficients when KLD and EM have been coupled (AIM # 3). Comparing the tables we observe that we are getting the same performance but the advantage is that we can perform inhomogeneity, feature selection and segmentation in one step. Similarly, Table 5.3 and Table 5.4 show the similarity coefficients for C vs. NT and C vs. T respectively.

Table 5.1 Summary of similarity coefficient for 8 patients using KLD as feature selection and EM as segmentation separately (AIM # 1) for T vs. NT.

Patient	Jaccard			Dice			Sokal & Sneath (SS)			Russel & Rao (RR)		
	T1	T2	Flair	T1	T2	Flair	T1	T2	Flair	T1	T2	Flair
1	0.62	0.65	0.61	0.71	0.73	0.73	0.47	0.49	0.53	0.62	0.64	0.6
2	0.71	0.69	0.73	0.84	0.75	0.78	0.78	0.65	0.65	0.71	0.7	0.72
3	0.72	0.69	0.68	0.78	0.76	0.8	0.7	0.67	0.66	0.75	0.71	0.65
4	0.8	0.81	0.72	0.83	0.88	0.84	0.78	0.77	0.7	0.8	0.82	0.72
5	0.8	0.68	0.83	0.84	0.74	0.85	0.76	0.63	0.8	0.81	0.68	0.83
6	0.82	0.85	0.81	0.86	0.9	0.85	0.72	0.75	0.73	0.84	0.84	0.81
7	0.77	0.86	0.79	0.8	0.88	0.83	0.74	0.84	0.75	0.78	0.87	0.8
8	0.84	0.87	0.83	0.87	0.89	0.88	0.8	0.84	0.79	0.82	0.85	0.82

Table 5.2 Summary of similarity coefficient for 8 patients using KLD as feature selection and EM as segmentation in integrated framework (AIM # 3) for T vs. NT.

Patient	Jaccard			Dice			Sokal & Sneath (SS)			Russel & Rao (RR)		
	T1	T2	Flair	T1	T2	Flair	T1	T2	Flair	T1	T2	Flair
1	0.65	0.67	0.61	0.72	0.74	0.73	0.52	0.51	0.53	0.63	0.64	0.64
2	0.71	0.69	0.73	0.85	0.75	0.78	0.78	0.65	0.65	0.71	0.73	0.72
3	0.73	0.69	0.68	0.78	0.76	0.84	0.7	0.67	0.66	0.75	0.71	0.65
4	0.8	0.81	0.72	0.83	0.88	0.84	0.78	0.77	0.7	0.84	0.82	0.72
5	0.8	0.68	0.83	0.84	0.74	0.85	0.76	0.63	0.8	0.81	0.68	0.83
6	0.82	0.85	0.81	0.86	0.9	0.85	0.72	0.75	0.73	0.84	0.84	0.81
7	0.77	0.86	0.79	0.8	0.88	0.83	0.74	0.84	0.75	0.78	0.87	0.83
8	0.84	0.88	0.83	0.87	0.89	0.88	0.8	0.84	0.79	0.82	0.85	0.82

Table 5.3 Summary of similarity coefficient for 8 patients using KLD as feature selection and EM as segmentation in integrated framework (AIM # 3) for C vs. NT.

Patient	Jaccard			Dice			Sokal & Sneath (SS)			Russel & Rao (RR)		
	T1	T2	Flair	T1	T2	Flair	T1	T2	Flair	T1	T2	Flair
1	0.65	0.67	0.61	0.62	0.72	0.63	0.72	0.65	0.67	0.63	0.64	0.64
2	0.70	0.69	0.70	0.65	0.65	0.68	0.68	0.65	0.65	0.71	0.63	0.72
3	0.63	0.69	0.64	0.68	0.66	0.68	0.71	0.67	0.68	0.75	0.71	0.65
4	0.62	0.61	0.65	0.68	0.67	0.64	0.67	0.71	0.74	0.64	0.65	0.72
5	0.63	0.71	0.62	0.64	0.73	0.73	0.72	0.65	0.72	0.71	0.68	0.63
6	0.62	0.62	0.62	0.66	0.71	0.71	0.71	0.73	0.68	0.74	0.64	0.68
7	0.65	0.70	0.64	0.64	0.68	0.73	0.65	0.68	0.69	0.72	0.67	0.63
8	0.64	0.69	0.62	0.68	0.68	0.69	0.68	0.69	0.65	0.62	0.65	0.62

Table 5.4 Summary of similarity coefficient for 8 patients using KLD as feature selection and EM as segmentation in integrated framework (AIM # 3) for C vs. T.

Patient	Jaccard			Dice			Sokal & Sneath (SS)			Russel & Rao (RR)		
	T1	T2	Flair	T1	T2	Flair	T1	T2	Flair	T1	T2	Flair
1	0.64	0.66	0.61	0.62	0.70	0.66	0.71	0.64	0.64	0.61	0.65	0.65
2	0.65	0.67	0.70	0.61	0.73	0.64	0.65	0.63	0.63	0.65	0.73	0.72
3	0.73	0.63	0.62	0.74	0.75	0.64	0.73	0.67	0.65	0.73	0.71	0.66
4	0.72	0.65	0.66	0.65	0.67	0.67	0.68	0.7	0.72	0.65	0.68	0.72
5	0.64	0.65	0.62	0.64	0.73	0.73	0.66	0.65	0.7	0.65	0.68	0.63
6	0.62	0.62	0.67	0.63	0.71	0.71	0.69	0.73	0.66	0.65	0.65	0.65
7	0.74	0.65	0.64	0.64	0.68	0.73	0.65	0.68	0.68	0.72	0.63	0.66
8	0.64	0.64	0.62	0.65	0.66	0.69	0.63	0.67	0.65	0.62	0.65	0.62

5.4. Conclusion

In this work we have coupled three steps – inhomogeneity correction, feature selection and segmentation in an EM framework. This framework selects the best feature for T and NT and performs segmentation simultaneously. This allows us to observe the effect of segmentation when selecting the best features. So we select a patient from the dataset, perform preprocessing, extract the texture features, and then feed the features to KLD-EM framework for selecting best features and segmentation. To obtain robustness

we use different similarity coefficients. We observe that the segmentation performance is same as that in AIM#1 but in this method can perform the steps in single framework.

6. Conclusion and Future Work

6.1. Discussion and Future Work

The primary goal of this dissertation is to investigate and improve robustness of feature –based pediatric PF tumor segmentation. The investigation is focused on selecting the best features from among multiple different features for tumor and non tumor. We implement an information theoretic approach for selecting the best features among multiple different features including out texture features. To improve the tumor segmentation we investigate features for abnormalities such as cyst. To achieve this goal we extend two class approaches to multiclass information theoretic. Finally, we present an integrated framework information theoretic approach for feature selection and segmentation in an EM framework.

6.2. Major Contributions

The major contribution of this dissertation includes three different novel computational models for improving the feature selection for effective segmentation of pediatric brain tumors. These computational models and associated studied are all published in proceeding of few major conferences [116], journal paper [69] an book chapter [134]. Now the contributions in these three dissertation are summarized as follows. In Chapter 3 discusses efficacy of texture, shape and intensity feature fusion for posterior-fossa tumor segmentation in MRI. The primary goal is to select the best feature for two class i.e. tumor and non tumor using KLD. For selection of the best feature, we compare four different techniques such as PCA, boosting, KLD and entropy metrics. We implement an integrated mathematical framework for feature selection and ranking using

KLD since KLD offers the best feature selection performance for the study. Our KLD feature selection technique shows that mBm is the best feature for both T1 and FLAIR modality while intensity is for T2 modality. In order to obtain robust segmentation of PF tumor in pediatric brain MRI, we compare performance of three different techniques such as bottom up top down, graph cut and EM. We finally select an integrated KLD - EM framework for tumor segmentation since this specific combination offers the best performance among the techniques investigated in this study. We evaluate robustness of our proposed model using four different similarity metrics and demonstrate the efficacy of our technique using 249 real MR images from ten pediatric patients. Furthermore, we show that fusion of mBm feature in multimodality T1, T2 and FLAIR MRI, can offer 100% PF tumor segmentation for the patient cases studied in this work.

In Chapter 4 we investigate information Theoretic Multiclass Feature Selection for Improved Pediatric Brain Tumor Segmentation. The goal is to select extend two class KLD to multiclass for selecting features for T, C and NT. We also improve feature selection by including abnormalities such as cyst as another class. We develop an information theoretic approach for multiclass feature selection for improved pediatirc brain tumor segmentation by combining KLD and the upper bound Baye's error. We obtain the best features which has minimun upper bound Baye's error from among set of features. Our Baye's KLD approach shows that mBm is the best feature for T / NT segmentation for T1 and FLAIR modalities.In addition, intensity is best for T / C, C / NT segmentation in T1, T2 and FLAIR modalities. In order to validate the segmentation efficacy, we obtain the pixel count in each of the T, C and NT regions using intensity and

mBm features. We compute segmentation robustness by comparing the tumor segments to the ground truth provided by radiologists for all eight pediatric patients.

Finally in Chapter 5 we develop an integrated framework for inhomogeneity, feature selection and segmentation in PF tumors. The goal is to integrate the three steps such as inhomogeneity correction, feature selection and segmentation in an EM framework. We develop a statistical framework using EM algorithm to couple these steps. This framework selects the best feature for two tissue type such as T and NT; T and C; and C and NT at a time for corresponding tissue segmentation. This allows us to observe the effect in segmentation at different iterations of inhomogeneity corrections and feature selection. We select a patient from the dataset, perform inhomogeneity preprocessing, extract the texture features, and then use the features to KLD-EM framework for selecting best features and subsequent segmentation in each iterations. We extensible validate robustness using different similarity coefficients. We observe that though the tumor tissue segmentation performance is similar to that in Chapter 3, however, we obtain such tumor segmentation performance in a single framework. Furthermore, our integrated framework allows one to observe the effect of inhomogeneity correction and feature selection on tumor segmentation step – by – step.

6.3. Future Works

In this section a few interesting future direction of this dissertation are discussed. In Chapter 3 we discuss features for T and NT. We need to investigate additional features for differentiating among tumor, non-tumor and edema. This will require work in extending KLD to discriminate multiclass tissues such as brain tissues, tumor, edema and

other artifacts in MRI. We also plan to investigate features for tumors such as astrocytoma, medulloblastoma, glioblastoma multiforme (GBM).

In this Chapter 4 we discuss relevant features for T, C and NT. We have used intensity as the features for discriminating C from other tissues. The segmentation can be improved further if we exploit other features depending on the characteristic of cyst for different tumors. We plan to develop other statistical models for feature selection.

In Chapter 5 we discuss integrated method for inhomogeneity, feature selection and segmentation. But we perform feature extraction offline after inhomogeneity correction. We resume feature selection and segmentation after feature extraction. We would like to come up with mathematical model for including feature extraction as the integrated step. This would add another parameter in the integrated framework and its effect on the existing parameters such as inhomogeneity, feature selection and segmentation.

REFERENCES

[1] Available at: http://www.mdanderson.org/children, May 2007.

[2] Available at: http://kidshealth.org/parent/medical/brain/brn_tumors.html, April 2005.

[3] A.A. Galhoun,H.H.Engelhard, and H.A.Al-Shatoury, " Posterior fossa tumors", *eMedicine Specialities*, available at: http://www.emedicine.com/med/topic2907.html, August 2006.

[4] C.T. Chin, S.Chang, and W.P.Dillon, "Brain and Spinal cord tumors," *Oncologic Imaging*, editor D.G. Bragg, 2nd edition, Missouri: Matthews Nook Company, pp. 133-159, 2002.

[5] K.M.Iftekharuddin, "Techniques in fractal analysis and their application in brain MRI," *Medical Imaging Systems Technology: Analysis and Computational Methods*, 2005.

[6] L.P.Clarke, R.P. Velthuizen, M.A. Camacho, J.J.Heine, M. Vaidyanath, L.O. Hall, R.W. Thatcher and M.L. Silbiger, "MRI Segmentation: methods and applications," *Magn. Resonance Imaging*, vol.13, no. 3, pp. 343-68, 1995.

[7] J.C. Bezdek, L.O. Hall and L.P. Clarke, "Review of MR image segmentation techniques using pattern recognition," *Medical Physics*, vol. 20, no. 4, pp. 1033-1048, 1993.

[8] H.S. Zadech and J.P. Windham, "Optimal linear transformation for MRI feature extraction enhancement and segmentation of magnetic resonance image scene sequence," *IEEE Trans. On medical Imaging*, vol. 11, pp. 302-318, 1992.

[9] M.E. Algorri and F. Flores – Mangas, "Classification of anatomical structures in MR brain images using fuzzy parameters," *IEEE Trans. Biomed Eng*, vol. 51, no. 9, pp. 1599-608, 2004.

[10] S. Warfield, J. Dengler, J. Zaers, C. Guttman, W. Wells, G. Ettinger, J. Hiller and R. Kikinis, "Automatic identification of gray matter structures from MRI to improve the segmentation of white matter lesions," *J. Image Guided Surgery*, vol. 1, pp. 326-338, 1995.

[11] S. Warfield, M. Kaus , F. Jolesz and R. Kikinis, "Adaptive template moderated spatially varuing statistical classifications," *Proc. of Int. Conf. on Medical Image Comp. and Comp. Asst. Inv.*, pp. 431-438, 1998.

[12] M. Prastawa, E. Bullitt, N. Moon, K. V. Leemput and G. Gerig, "Automatic brain tumor segmentation by subject specific modification on atlas priors," *Acad. Radio*, vol.10, pp. 1341 – 1348, 2003.

[13] Y. Li and Z. Chi, "MR Brain Image Segmentation Based on Self-Organizing Map Network," *International Journal of Information Technology,* vol. 11, no. 8, 2005.

[14] J. Zhang, K.Ma and M. Er, "Tumor Segmentation from Magnetic Resonance Imaging by Learning via one class Support Vector machine," 2004.

[15] D.T. Gering, "Recognizing Deviations from Normalcy for Brain Tumor Segmentation," PhD. Thesis, MIT, 2003.

[16] S. Herlidou, Y. Rolland, J.Y. Bansard, E.L. Rumeur and J.D. Certains, "Comparison of automated and visual texture analysis in MRI: characterization of

normal and diseased skeletal muscle," *Magnetic Resonance. Imaging*, vol. 19, no.9, pp. 1393-1397, 1999.

[17] D. Mahmoud – Ghoneim, J.D. Certains, S. Herlidou, Y. Rolland and A. Maniere, "A gender difference on magnetic resonance imaging texture analysis on human adipose tissue," *Journal of Women Imag*ing, vol. 3, pp. 105-107, 2001.

[18] P.A. Eliat, D. Mahmoud – Ghoneim and J.D. Certains, "Texture analysis and metabolic magnetic resonance imaging," *Trends Appl. Spect*., vol. 4, pp. 1-20, 2002.

[19] R. A. Lerski, K. Straughan, L.R. Schad, D. Boyce, S. Bluml and I. Zuna, "MR image texture analysis: an approach to tissue characterization," *Magnetic Resonance Imaging*, vol. 11, pp. 873-887, 1993.

[20] D. Mahmoud – Ghoneim, G. Toussiant, J. Constans and J.D. Certains, "Three dimensional texture analysis in MRI: a preliminary evalutaionn in gliomas," *Magnetic Resonance Imaging*, vol. 21, pp. 983-987, 2003.

[21] C. Pachai, Y.M. Zhu, J. Grimand, M. Hermier, A. Dromigny – Badin, A. boudraa, G. Giemenez, C. Confavreux and J.C. Froment, "A pyramidal approach for automatic segmentation of multiple sclerosis lesions in brain MRI," *Computational medical Imaging Graph*, vol. 22, no. 5, pp. 399-408, 1998.

[22] A. Pitioit, A. Toga, N. Avache and P. Thomson, "Texture based MRI segmentation with a two stage hybrid neural classifier," *IEEE Int. Joint Conf. on Neural Network*, 2002.

[23] A. W.C. Liew and H. Yan, "Current Methods in the Automatic Tissue Segmentation of 3D Magnetic Resonance Brain Images," *Current Medical Imaging Reviews*, vol. 2, no. 1, 2006.

[24] X. Zeng, L.H. Staib, R.T. Schultz and J.S. Duncan, "Segmentation and measurement of the cortex from 3 D MR images using coupled surface propagation," *IEEE Trans. On Medical Imaging*, vol. 18, no.10, pp. 927-937, 1999.

[25] N. Sarkar and B.B. Chaudhuri, "An efficient approach to estimate fractal dimension of textural images," *Pattern Recognition*, vol. 23: pp. 1035- 1041, 1992.

[26] P.M Thompson, A.D. Lee, R.A. Dutton, J.A. Geaga, K.M. Hayashi, M.A. Eckert, U. Bellugi, A.M. Galaburda, J.R. Korenberg, D.L. Mills, A.W. Toga and A.L. Reiss, "Abnormal cortical complexity and thickness profiles mapped in Williams syndrome," *Journal of Neuroscience*, vol.25, no.16, pp. 4146-58, 2005.

[27] A.Islam, K. M. Iftekharuddin, R.Ogg, F.H. Laningham and B. Sivakumar, " Multifractal modeling, segmentation, prediction and statistical validation of posterior fossa tumor," *SPIE Medical Imaging Conference*, Feb2008

[28] J. M. Zook and K. M. Iftekharuddin, "Statistical analysis of fractal-based brain tumor detection algorithms," *Magnetic Resonance Imaging*, vol. 23, pp. 671-678, 2005.

[29] K.M. Iftekharuddin, A. Islam, J. Shaik, C. Parra and R. Ogg, "Automatic brain tumor detection in MRI: methodology and statistical validation," *Proceedings of the SPIE Symposium and Medical Imaging*, vol. 5747, pp. 2012-2022, 2005.

[30] K.M. Iftekharuddin, J.Zheng, M.A. Islam and R.J.Ogg, “Fractal –based Brain Tumor Detection in Multimodal MRI,” *Journal of Applied Mathematics and Computation*, vol. 206, 2009.

[31] K.M. Iftekharuddin, W. Jia and R. Marsh, “A fractal analysis of tumor in brain MR images,” *Machine vision application*, vol. 13, pp. 352-262, 2003.

[32] K.M. Iftekharuddin and C. Parra, “Improved multiresolution – fractal feature extraction for tumor detection in MR images,” *2nd Int. Cong. Of Computational Harmonic Analysis*, May 2004.

[33] S. Bidiwala and T. Pittman, “Neural Network Classification of Pediatric Posterior Fossa Tumors Using Clinical and Imaging Data,” *Journal of Pediatric Neurosurgery*, vol.40, pp. 8-15, 2004.

[34] K. Xie , J. Yang, Z.G. Zhang and Y.M. Zhu, “ Semi automated brain tumor and edema segmentation in MRI,” *European Journal of Radiology*, vol. 56, no.1, pp. 12-19, 2005.

[35] G.Z. Li, J. yang, C.Z. Ye and D. Geng, “Degree prediction of malignancy in brain gliomas using support vector machines,” *Computers in Biology and Medicine*, 2005.

[36] T. Dietterich, “Machine learning research: Four current directions,” *Journal of Artificial Intelligence*, vol.18, pp. 97-136, 1998.

[37] H. Chouaib, O.R. Terrades, S. Tabbone, F. Cloppet and N. Vincent, “Feature selection combining genetic algorithm and adaboost classifiers,” *IEEE transaction on Pattern Recognition*, vol. 8, no.11, pp. 1 - 4, 2008.

[38] G.D. Tourassi, E. D. Frederick, M.K. Markey and C. E. Floyd, “Application of mutual information criterion for feature selection in computer diagnosis,” *American Association of Physicists in Medicine*, 2001.

[39] Du Zheng , Y.S. Jeong, M.K. Jeong and S.G. Kong, “Multidimensional local spatial autocorrelation measure for integrating spatial and spectral information in hyperspectral image band selection,” Journal in Applied Intelligence, 2011.

[40] K.M. Pohl, J. Fisher, J.J. Levitt, M.E. Shenton, R. Kikinis, W.E.Grimson and W.M. Wells, “A unifying approach to registration, segmentation and intensity correction,” *MICCAI* 2005.

[41] M.N. Ahmed, S.M. Yamany, A.A. Farag and T. Moriarty, “A modified C-means algorithm for bias field estimation and segmentation of MRI data,” *IEEE transaction on Medical Imaging*, vol.21, no. 3 , March 2002.

[42] L. Lihong, W. Xinzhou and L. Zhengrong, “A unifying framework for inhomogeneity correction and partial volume segmentation of brain MR images,” *MICCAI*, 2004.

[43] S. Warfield, M. Kaus, F.A. Jolesz and R. Kikinis, “Adaptive template moderated spatially varying statistical classification,” *MICCAI*, vol.1496, pp. 431-438, Springer, 1998.

[44] S. Warfield, J. Dengler, J. Zaers, C.R.G. Guttman,W.M.Wells, G.J. Ettinger, J.Hiller and R. Kikinis, “Automatic identification of gray matter structures from MRI to improve the segmentation of white matter lesions,” *MICCAI*, vol.1, no.6, pp. 326-338, 1996.

[45] A. Bru and J.M. Pastor, "Super –rough dynamics on tumor growth," *Physical Review Letters*, vol. 81, no.18, pp.4008-4011, 1998.

[46] J. Yu and Y. Wang, "Molecular Image Segmentation based on Improved Fuzzy Clustering", *Int. Journal of Biomedical Imaging*, vol. 20, 2007.

[47] V. Raad, "Design of Gabor wavelets for analysis of texture features in cervical imaging," *Proceedings of 23th Annual Int. Conf. of IEEE*, vol.1, pp.17-21, 2003.

[48] Y. Zhan and D. Shen, "Deformable Segmentation of 3D Ultrasound Prostrate Images using Statistical Texture Matching method," *IEEE transaction of Medical Imaging*, vol. 25, no.3, 2006.

[49] C. Parra and K. M. Iftekharuddin, "Multiresolution-fractal feature extraction and tumor detection: analytical modeling and implementation," *Proceedings of the 47th Annual SPIE Meeting, Optical Science and Technology*, 2003.

[50] S. Bidiwala and T. Pittman, "Neural Network Classification of Pediatric Posterior Fossa Tumors Using Clinical and Imaging Data," *Journal of Pediatric Neurosurgery*, vol.40, pp. 8-15, 2004.

[51] X. Xu and A. Zhang, "Boost Feature Subset Selection: A new gene selection algorithm for Microarray Dataset," *International Journal of Data Mining and Bioinformatics*, vol. 3, 2009.

[52] I. Cohen, Q. Tan, X.S. Zhou and T.S. Huang, "Feature Selection using Principal Feature Analysis," *ACM Multimedia*, 2007.

[53] J. Novovicova, P. Pudil and J. Kittler, "Divergence Based Feature Selection for Multimodal Class Densities," *IEEE transaction on Pattern Analysis and Machine Intelligence*, vol.18, no.2, 1996.

[54] E. Borenstein, E. Sharon and S. Ullman, "Combining Top down and Bottom up segmentation," *Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop*, vol. 4, 2004.

[55] M. Schaap, L. Neefjes, C. Metz, A. Giessen, A. Weustink, N. Mollet, J. Wentzel, T. Walsum and W. Niessen, "Coronary Lumen Segmentation using graph cuts and robust kernel regression," *Information Processing in Medical Imaging*, 2009.

[56] S. Borman,"The Expectation Maximization Algorithm-A Short Tutorial," available at *http://www.seanborman.com/publications/EM_algorithm.pdf*, 2004.

[57] B.B Mandelbrot, "The fractal geometry of nature," *San Francisco;Freeman*, 1983.

[58] P. Goncalves, "Existence test of moments: Application to Multifractal Analysis," Proceedings Int. Conf. of telecommunications, 2000.

[59] R. Malladi, J. A. Sethian and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE transaction on Pattern analysis and Machine intelligence*, vol. 17, no.2, 1995.

[60] S. Suri ,K. Liu, S. Singh , S. N. Laxminarayan, X. Zeng and L. Reden , "Shape Recovery Algorithms Using Level Sets in 2-D/3-D Medical Imagery: A State-of-the-Art Review," *IEEE transaction on information technology in Biomedical*, vol.6, no.1, 2002.

[61] G. Zhu, S. Zhang, Q. Zeng and C. Wang, "Boundary – based image segmentation using binary level set method," Opt.Eng, vol.46, 2007.

[62] T. Kavzoglu and P.M. Mather, "The use of feature selection techniques in the context of artificial neural networks," *Proceedings on 26th Annual Conference of Remote Sensing Society*, 2000.

[63] J. Zhang and R. J. Figueiredo, "Autonomic Feature Selection for Application Classification," *IEEE conference on Autonomic computing*, 2006.

[64] G. Xuan, Z. Zhang, P. Chai, Y.Q. Shi and D. Fu, "A Feature Selection Based on Minimum Upper Bound of Baye's Error," *IEEE workshop on Multimedia Signal Processing*, vol. 2, pp.1-4, 2006.

[65] L. Bruzzone and S.B. Serpico, "A technique for feature selection in multiclass problems," *Int. Journal on Remote Sensing*, vol.21, no. 3, pp. 549-563, 2000.

[66] L. Bruzzone, F. Roli and S.B. Serpico, "An extension of the Jeffreys – Matusita Distance to Multiclass cases for Feature Selection," *IEEE transaction on Geoscience and Remote Sensing*, vol.33, no.6, 1995.

[67] L.Bruzzone, D.F.Prieto and S.B.Serpico, "A neural – statistical approach to multitemporal and multisource remote - sensing image classification," *IEEE transaction on Geoscience and Remote Sensing*, vol. 37, no. 3, 1999.

[68] L.Bruzzone and S.B.Serpico, "An iterative technique for the detection of land cover transitions in multitemporal remote - sensing images," *IEEE transaction on Geoscience and Remote Sensing*, vol. 35, no. 4, 1997.

[69] S. Ahmed, K.M. Iftekharuddin and A. Vossough, "Efficacy of texture, shape and intensity feature fusion for robust posterior – fossa tumor segmentation in MRI," *IEEE transaction for Information Technology in Medicine and Biology Society*, vol. 15, no. 2. pp. 206-213, 2011.

[70] S. Ahmed, K.M. Iftekharuddin, R. J. Ogg and F. H. Laningham, "Efficacy of texture, shape and intensity features for robust posterior – fossa tumor segmentation in MRI," *Proceeding of SPIE Medical Imaging*, 2008.

[71] K.L. Lee and L.H. Chen, "Unsupervised texture segmentation by determining the interior of texture regions based on wavelet transform," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 15, no. 8, pp. 1231- 1250, 2001.

[72] H. Greenspan, A. Ruf and J. Goldberger, "Constrained Gaussian mixture model framework for automatic segmentation of MR brain images," *IEEE transaction on medical imaging*, vol. 29, no. 9, 2006.

[73] K. V. Leemput, F. Maes, D. Vandermeulen and P. Suetens, "Automated model-based tissue classification of MR images of the brain," *IEEE Trans. Med. Imag*., vol. 18, no. 10, pp. 897–908, 1999.

[74] J. L. Marroquin, B. C. Vemuri, S. Botello, F. Calderon and A. Fernandez-Bouzas, "An accurate and efficient bayesian method for automatic segmentation of brain MRI," *IEEE Trans. Med. Imaging*, vol. 21, no. 8, pp. 934–945, 2002.

[75] M. Prastawa, J. Gilmore, W. Lin and G. Gerig, "Automatic segmentation of neonatal brain MRI," *Proc. Int. Conf. Med. Image Comput. Comput. Assisted Intervention (MICCAI)*, pp. 10–17, 2004.

[76] A.Riad, A. Atwan, H.K. Elminir and N.Mastorakis, " A new approach for segmentation of brain MR image," Proceedings of WSEAS international conf. on environment, medicine and health science, 2006.

[77] W.M. Wells, W.E.L. Grimson, R. Kikinis and F.A. Jolesz, “Statistical gain correction and segmentation of MRI data,” Draft of August 5, internal report BWH Boston, 1993.

[78] William Wells, Ron Kikinis and Ferenc A. Jolesz, “Statistical intensity correction and segmentation of magnetic resonance image data,” *Proceedings of the Third Conference on Visualization in Biomedical Computing*, vol. 2359, pp. 13-24, SPIE, 1994.

[79] W.M. Wells, W.E.L. Grimson, R. Kikinis and F.A. Jolesz, “Adaptive Segmentation of MRI Data”, *IEEE Transactions on Medical Imaging*, vol. 15, no. 4, pp. 429 - 443,1996.

[80] R. Guillemaud and M. Brady, “Estimating the Bias Field of MR Images,” *IEEE Transactions on Medical Imaging*, vol. 16, no. 3, pp. 238 - 251, 1997.

[81] K. Van Leemput, F. Maes, D. Vandermeulen and P. Suetens, “Automated model-based tissue classification of MR images of the brain,” *IEEE TMI*, Vol. 18, pp. 897- 908, 1999.

[82] K. Van Leemput, F. Maes, D. Vandermeulen and P. Suetens, “Automated model-based bias field correction of MR images of the brain,” *IEEE TMI*, vol. 18, pp. 885- 896, 1999.

[83] D.T. Gering, W.E.L. Grimson and R. Kikinis, “Recognizing deviations from normalcy for brain tumor segmentation,” LNCS 2488, Proceedings *in MICCAI*, pp. 388-395, 2002.

[84] K. Pohl, S. Bouix, R. Kikinis and W. Grimson, "Anatomical guided segmentation with non-stationary tissue class distributions in an expectation-maximization framework," in *ISBI*, pp. 81–84, 2004.

[85] K.Poh, J.Fischer, J.J.Levitt, M.E. Shenton, R. Kikinis, W.E.Grimson and W. Wells, "A unifying approach to registration, segmentation and intensity correction," *Eighth International Conference on Medical Image Computing and Computer Assisted Intervention*, Springer-Verlag, LNCS 3749, pp. 310-318, 2005.

[86] A. W.C. Liew and H. Yan, "Current Methods in the Automatic Tissue Segmentation of 3D Magnetic Resonance Brain Images," *Current Medical Imaging Reviews*, vol. 2, no. 1, 2006.

[87] N. Sarkar and B.B. Chaudhuri , "An efficient approach to estimate fractal dimension of textural images," *Pattern Recognition*, vol. 23, pp. 1035- 1041, 1992.

[88] P.M Thompson, A.D. Lee, R.A. Dutton, J.A. Geaga, K.M. Hayashi, M.A. Eckert, U. Bellugi, A.M. Galaburda, J.R. Korenberg, D.L. Mills, A.W. Toga and A.L. Reiss, "Abnormal cortical complexity and thickness profiles mapped in Williams syndrome," *Journal of Neuroscience*, vol. 25, no.16, pp. 4146-58. 2005.

[89] S. Osher and J.A. Sethian, "Fronts propagating with curvature dependent speed: algorithms based on Hamilton – Jacobi formulation," *Journal of Computational Physics*, vol. 79, pp. 12-49, 1988.

[90] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A Shape-Based Approach to the Segmentation of Medical

Imagery Using Level Sets," *IEEE Transaction on Medical Imaging*, vol. 22, no. 2, 2003.

[91] Y.Sun, C.F. Babbs and E.J. Delp, "A comparison of feature selection methods for the detection of breast cancer in mammograms: adaptive sequential floating search vs. genetic algorithm," *IEEE-EMBS*, pp.6532-6535, 2005.

[92] M.F. McNitt - Gray, H.K. Huang and J.W. Sayre, " Feature selection in the pattern classification problem of digital chest radiograph segmentation," *IEEE Transactions* on *Medical Imaging,* vol. 14, no. 3, pp.537-547,1995.

[93] N.I. Weisenfield and S.K.Warfield, "Normalization of Joint Image –Intensity Statistics in MRI using the Kullback – Leibler Divergence," *IEEE International Symposium on Biomedical Imaging: Nano to Macro,* vol.1, no.15-18, pp.101-104, 2004.

[94] Zhu C.M., GU Guo-Chang, L.Hai-Bo, S. Jing and Y. Hualong, "Segmentation of Ultrasound Image Based on Tetxure Feature and Graph Cut," *International Conf. on Computer Science & Software Engg*., 2008.

[95] J. Shi, J.Malik, "Normalized Cuts and Image Segmentation," *IEEE Trans. On Pattern Analysis & Machine Learning*, pp: 888-905, 2000.

[96] P. Tsui, "Expectation Maximization Algorithm for Gaussian Mixture Model," available at: *http://www.mathworks.com/matlabcentral/fileexchange/*, 2006.

[97] A.G. Osborn and M. T. Greece, "Intracranial cysts: Radiologic – Pathologic Correlation and Imaging Approach," *Journal of American Radiology*, vol.3, pp. 650-664, 2006.

[98] Q. Wang, E. K. Liacouras, E. Miranda ,U. S. Kanamalla and V. Megalooikonomou , "Classification of brain tumors in MR images," *Int. Society for optical engineering*, vol. 6514, no.1, pp.6514-32,1995.

[99] P. Golland, W.E.Grimson, M.E. Shenton and R. Kikinis, "Detection and analysis of statistical differences in anatomical shape," *Med Image Analysis*, vol. 9, no.1, pp.69-86, 2005.

[100] S. Zafeiriou , A. Tefas and I. Pitas, "Linear Discriminant feature selection techniques in elastic graph matching," *Journal of Pattern Recognition*, vol.40, no. 10, 2007.

[101] Y. Yu , J. Huang, S. Zhang , C. Restif , X. Huang , D. Metaxas, "Group sparsity based classification for cervigram segmentation," ISBI11.

[102] K. Xiao, S.H.Ho and A. Bargiela, "Automatic brain MRI segmentation scheme based on feature weighting factors selection on fuzzy c-means clustering algorithms with Gaussian smoothing," *Int. Journal in Computational Intelligence in Bioinformatics and System Biology*, vol.1, no. 3, 2010.

[103] G. Li, J. Yang, G. Liu and Li Xue, "Feature selection for multiclass problems using support vector machine," *Lecture Notes in Computer Science*, 3157 Springer 2004, ISBN 3-540-22817-9.

[104] T. Li, Chengliang Zhang and Mitsunori Ogihara, "A comparative study of feature selection and multiclass classification methods for tissue classification based on gene expression," *Journal of Bioinformatics*, vol. 20, no.15, pp. 2429–2437, 2004.

[105] H. Lu, W. Chen, X. Ma, M. Wang and J. Zhang, "Model – free gene selection using Genetic Algorithm," *International Journal of Digital Content technology and its Application*, vol.5, no. 1, 2011.

[106] P. Daniel Berrar, C. S. Downes, Wernerdubitzky, "Multiclass cancer classification using gene expression profiling and Probabilistic Neural Network," *In Proceedings of the Pacific Symposium on Biocomputing*, 2003

[107] X. Chen, X. Zheng and D. Alphen, "Multiclass feature selection for texture classification," *Journal of Pattern recognition*, vol. 27, no.14, 2006.

[108] D. Joeng, D. Kang and S. Won, "Feature selection for multiclass classification using support vector data description," *36th Annual Conference on IEEE Industrial Electronics Society*, pp. 1129 – 1132, 2010.

[109] J.Yi Lin, H.Ren Ke , B.Chian Chien and W.Pang Yang, "Classifier design with feature selection and feature extraction using layered genetic algorithm," *Expert Systems with Application*, vol.34, pp. 1384–1393, 2008.

[110] L. Rokach, "Genetic algorithm based feature set partitioning for classification problems," *Journal of Pattern Recognition*, vol. 41, 2008.

[111] H. Vafaie and I.F. Imama, "Feature selection methods: Genetic algorithms vs. Greedy like search," *Journal of Artificial Intelligence*, vol. 151, no.1, pp.155-176, 2003.

[112] A. Verikas and M. Bacauskiene, "Feature selection with neural networks," *Journal of Pattern Recognition*, vol. 23, no. 11, 2002.

[113] L. Backstrom and R. Caruana, "C2FS: An algorithm for feature selection in Cascade neural network," *IJCNN*, 2006.

[114] W. Chen, R. M. Zur and M.L. Giger, “Joint feature selection and classification using a Bayesiam neural network with automatic relevance determination priors: potential use in CAD of medical imaging,” *Proceedings of Computer Aided Diagnosis in Medical Imaging*, vol. 6514, 2007.

[115] H. Yuan, S.S. Tseng, Wu Gangshan and Z. Fuyan, “A two phase feature selection method using both filter and wrapper,” *IEEE International Conference on Systems, Man, and Cybernetics*, 1999.

[116] S. Ahmed, K.M. Iftekharuddin, R. J. Ogg and F. H. Laningham, “Efficacy of texture, shape and intensity features for robust posterior – fossa tumor segmentation in MRI,” *Proceeding of SPIE Medical Imaging*, 2008.

[117] A.Simmons, P.S.Tofts, G.J.Barker and S.R.Arridge , “Sources of intensity nonuniformity in spin echo images at 1.5T ,” *Magnetic Resonance in Medicine*, vol. 32, pp.121-128, 1994.

[118] J.G. Sled and G.B. Pike, “Standing wave and RF penetration artifacts caused by elliptic geometry: an electrodynamic analysis of MRI,” *IEEE Transactions on Medical Imaging*, vol. 1, pp. 653-662, 1998.

[119] J.A.Weszka, “A survey of threshold selection techniques,” *Computer Graphics and Image processing*, vol. 7, pp. 259-265, 1978.

[120] R.B.Lufkin, T. Sharpless, B. Flanningan and W. Hanafee, “Dynamic range compression in surface coil MRI,” *AJR*, vol. 147, no.379, pp. 379-382, 1986.

[121] L. Axel, J. Costantini and J. Listerud, “Intensity correction in surface coil MR imaging,” *AJR*, vol. 148, no.4, pp. 418-420, 1987.

[122] K.O.Lim and A. Pfferbaum, “Segmentation of MR brain imaging into cerebrospinal fluid spaces, white and gray matter,” *JCAT*, vol.13, no.4, pp. 588-593, 1989.

[123] B. Dawant, A. Zijdenbos and R. Margolin, “Correction of Intensity variations in MR images for computer aided tissue classification,” *IEEE transaction of medical imaging*, vol. 12, no.4, pp. 770 – 781, 1993.

[124] J. Gohagan, E. Spitzznagel, W. Murphy and M. Vannier, “Multispectral Analysis of MR images of breast,” *Radiology*, vol.163, pp. 703-707, 1987.

[125] M. Styner, C. Brechiihler, G. Szekely and G.Gerig, “Parametric estimate of intensity inhomogeneities applied to MRI,” *IEEE transaction on Medical Imaging*, vol. 19, no. 3, pp. 153 -165, 2000.

[126] C. Li, R. Huang, Z. Ding, C. Gatenby and D. Mataxas, “A variational level set approach to segmentation and bias correction of images with intensity in homogeneity,” *MICCAI*, vol. 11, no. 2, pp. 1083 – 1091, 2008.

[127] D. Pham and J. Prince, “Adaptive fuzzy segmentation of magnetic resonance images,” *IEEE Trans. Medical Imaging*, vol. 18, no.9, pp.737–752, 1999.

[128] M.I. Kohn, N.K. Tanna, G.T. Herman, S.M. Resnick, P.D. Mozley, R.E. Gur, A. Alavi, R.A. Zimmerman and R.C. Gur, “Analysis of brain and cerebrospinal fluid volumes with MR imaging - Part I. Methods, reliability, and validation,” *Radiology*, vol. 178, pp. 115-122, 1991.

[129] M. Tincher, C.R. Meyer, R. Gupta and D.M. Williams, “Polynomial modeling and reduction of RF body coil spatial inhomogeneity in MRI,” *IEEE Transactions on Medical Imaging*, vol. 12, no. 2, pp. 361- 365, 1993.

[130] Charles R. Meyer, Peyton H. B land and James Pipe, “Retrospective correction of Intensity Inhomogeneities in MRI,” *IEEE Transactions on Medical Imaging*, vol. 14, no. 1, pp. 36- 41, 1995.

[131] S. Warfield, J. Dengler, J. Zaers, C.R.G. Guttman, W.M. Wells, G.J. Ettinger, J. Hiller and R. Kikinis, “Automatic identification of gray matter structures from MRI to improve the segmentation of white matter lesions,” *MICCAI*, vol. 1, no. 6, pp. 326 - 338, 1996.

[132] S. Warfield, M. Kaus, F.A. Jolesz, and R. Kikinis, “Adaptive template moderated spatially varying statistical classification,” *MICCAI*, pp. 431- 438, Springer, 1998.

[133] S. Bricq, Ch. Collet and J. P. Armspach, “Unifying framework for multimodal brain MRI segmentation based on Hidden Markov chain,” *Medical Image Analysis*, vol. 12, no. 6, pp. 639-652, 2008.

[134] S. Ahmed and K.M. Iftekharuddin, “Efficacy of texture features for Brain tumor segmentation,” to be published in Tumor Classification Book Volumes.