ORIGINAL ARTICLE

3D variational brain tumor segmentation using Dirichlet priors on a clustered feature set

Karteek Popuri · Dana Cobzas · Albert Murtha · Martin Jägersand

Abstract

Purpose Brain tumor segmentation is a required step before any radiation treatment or surgery. When performed manually, segmentation is time consuming and prone to human errors. Therefore, there have been significant efforts to automate the process. But, automatic tumor segmentation from MRI data is a particularly challenging task. Tumors have a large diversity in shape and appearance with intensities overlapping the normal brain tissues. In addition, an expanding tumor can also deflect and deform nearby tissue. In our work, we propose an automatic brain tumor segmentation method that addresses these last two difficult problems.

Methods We use the available MRI modalities (T1, T1c, T2) and their texture characteristics to construct a multi-dimensional feature set. Then, we extract clusters which provide a compact representation of the essential information in these features. The main idea in this work is to incorporate these clustered features into the 3D variational segmentation framework. In contrast to previous variational approaches, we propose a segmentation method that evolves the contour in a supervised fashion. The segmentation boundary is driven by the learned region statistics in the cluster space. We incorporate prior knowledge about the normal brain tissue appearance during the estimation of these region statistics. In particular, we use a Dirichlet prior that discourages the clusters from the normal brain region to be in the tumor region. This leads to a better disambiguation of the tumor from brain tissue.

Results We evaluated the performance of our automatic segmentation method on 15 real MRI scans of brain tumor patients, with tumors that are inhomogeneous in appearance, small in size and in proximity to the major structures in the brain. Validation with the expert segmentation labels yielded encouraging results: Jaccard (58%), Precision (81%), Recall (67%), Hausdorff distance (24 mm).

Conclusions Using priors on the brain/tumor appearance, our proposed automatic 3D variational segmentation method was able to better disambiguate the tumor from the surrounding tissue.

Introduction

Motivation

In the diagnosis of brain tumors, radiation oncologists extensively use the magnetic resonance image (MRI) imaging modality. The goal of the radiation treatment is to kill the whole tumor region while leaving the healthy brain intact. The effectiveness of this radiation treatment therefore depends on an accurate MRI segmentation, meaning correct labeling of the regions in the MRI image as tumor or healthy. In current practice, the expert radiation oncologist has to manually perform this tedious and highly time-consuming segmentation task. In addition to being extremely time consuming and tedious, manual segmentation has two other major drawbacks:

(1) There exists significant variation between the segmentation labels produced by different medical experts.

K. Popuri (✉) · D. Cobzas · M. Jägersand
Department of Computing Science, University of Alberta,
Edmonton, Canada
e-mail: kpopuri@ualberta.ca

A. Murtha
Department of Oncology, University of Alberta, Edmonton, Canada

Further, the segmentations produced by the same expert in different settings are also subject to variations [21].

(2) In most cases, the 2D MRI image slices are labeled independently without taking into account the global 3D brain structure, leading to potentially inaccurate segmentations. Subsequently, there have been a lot of efforts to develop semi-automatic and fully automatic segmentation algorithms to delineate tumors in MRI images.

The pathological process responsible for the creation and growth of brain tumors is inherently unpredictable. Consequently, the occurrence of tumors cannot be spatially localized to a particular region in the brain. Further, the geometric properties of the tumor do not conform to a particular shape/size distribution. This lack of availability of spatial or shape priors on tumors makes automatic segmentation more challenging compared to the segmentation of objects in natural images. The observed intensity distribution in the tumor region spans a wide range; therefore, the voxel intensities in the tumor regions often overlap with the normal tissues. This severely affects the performance of segmentation methods that solely depend on intensity information to discriminate between normal and tumor tissues [11]. Another aspect that complicates segmentation is that the tumors often push around the surrounding normal structures and in effect distort them. This makes it impossible to use any kind of shape prior on these normal structures to aid in the tumor segmentation. To address these challenges, we propose to augment the intensity information in the brain MRI image with a set of features that capture the texture and symmetry properties of the brain image, thereby providing our segmentation algorithm a better discriminatory power at the voxel level. Further, we propose to incorporate prior knowledge about the normal brain tissue appearance (instead of shape) into our segmentation method.

Related work

Most of the previous work on brain tumor segmentation was based on machine learning classification techniques, either unsupervised (clustering or fuzzy clustering) [2,10,34,37] or supervised [32,35,38] methods. A common drawback among these classification-based segmentation approaches is that they consider the voxels in the image to be independent of each other, having no spatial correlation both in the training and testing phases. The atlas-based tumor segmentation approaches [13,23,27] address this drawback by incorporating spatial information into the segmentation process using an atlas. However, as the atlases are built using brain MRI data from healthy individuals, in the case of MRI data from patients with brain tumors, errors might occur during the registration step of these atlas-based tumor segmentation methods. This is because the growth of tumors

tends to deform the normal structures in a patient's brain and hence the normal structures in a patient's MRI brain image might not correspond to the normal structures in an atlas. A more direct approach is to integrate the spatial correlation between the voxels into the segmentation objective itself. In the past few years, various energy optimization approaches with both discrete (graph based or random field based) [1,7,16,17] and continuous (variational or active contour based) [6,9,28,36] formulations have been explored with considerable interest for the task of tumor segmentation. These approaches formulate the problem of tumor segmentation as an optimization task where the segmentation boundary is obtained by the optimization of an appropriate energy function.

In a discrete approach to image segmentation, the image is modeled as an undirected graph. The nodes of the graph represent the pixels (or voxels) in the image, and the graph edges represent the neighborhood relationship between the pixels (or voxels). The popular random field methods perform image segmentation by minimizing an energy function defined on the graph corresponding to the image. Variants of such random field methods have been employed for the task of MRI brain tumor segmentation: Markov Random Field [3], Support Vector Random Field [16], Decoupled Conditional Random Field [17]. Other discrete segmentation approaches use graph-based image "clustering," like the normalized cuts [29]. The basic idea in these methods is to divide the image clusters purely based on the edge affinities between the neighboring pixels. Recently, such methods have also been explored for MRI brain tumor segmentation with moderate success [1,7].

In continuous approaches like the variational or active contour-based segmentation, starting with an initialization, the segmentation boundary is evolved based on the minimization of an energy functional until the desired segmentation is achieved. In general, the energy functional consists of two terms: the data term incorporating the image-based forces to drive the segmentation boundary and a regularization term which enforces the spatial smoothness constraint on the evolving segmentation boundary. Further, the level set framework is often employed, wherein the segmentation boundary is implicitly defined as the zero-level curve (surface in 3D) of an embedding function ϕ. The resulting evolution equation usually takes the following form:

$$\frac{\partial \phi}{\partial t} = (F_I + \nu F_\phi)|\nabla \phi| \tag{1}$$

where, F_I is the evolution speed dependent on the image information, F_ϕ is the speed due to the smoothness constraint and ν is the relative weighting between the data and smoothness terms.

The level set method is an ideal tool for the segmentation of anatomical structures in medical images. This is due to its

ability to handle complex geometries and arbitrary topological changes. Subsequently, there have been efforts to develop tumor segmentation methods using the level set framework. Droske et al. [9] presented a level set-based glioma (tumors with highly irregular shapes) segmentation method. In this method, the tumor boundary was determined from the edge information in the MRI image. The evolution speed function, F_I, was such that it would take very small values in the presence of an image gradient (edges). Further, this method implemented a fast multigrid strategy for the propagation of the segmentation boundary. In spite of the efficient implementation, as this method was purely edge based, it suffered from the problem of getting trapped into local minima. Consequently, this method required a substantial amount of user input, for both the parameter tuning of the edge detection function F_I and the initialization of the segmentation boundary, which is undesirable in a practical tumor segmentation setting.

In the influential work of Ho et al. [28], the authors proposed a level set method based on region competition for the segmentation of brain tumors. In this method, the segmentation boundary was guided by a global region-based force instead of local image gradients. This method was therefore successful in the detection of smooth tumor boundaries and was also more robust to initialization. The region-based force at each voxel was determined by a probability map $p_{\text{tumor}}(\mathbf{x}) - p_{\text{non-tumor}}(\mathbf{x})$. These tumor/non-tumor probabilities were estimated from a mixture density fitted to the histogram of the T1–T1c difference image (T1c is T1 with gadolinium contrast agent). The segmentation boundary was initialized automatically as the zero-level set of the probability map, and the evolution proceeded with speed $F_I \propto (p_{\text{tumor}}(\mathbf{x}) - p_{\text{non-tumor}}(\mathbf{x}))$ until it converged to the tumor boundary. A limitation of this method was that it could only be applied to the case of tumors with "enhancing" regions, which are visible in the T1–T1c difference image. The incorporation of other MRI modalities like T2, FLAIR into this methodology is not obvious. Xie et al. [36] proposed the hybrid level set (HLS) method, where both region- and edge-based terms are employed in an active contour framework for the segmentation of brain tumors and edema. Here, the segmentation boundary is primarily driven by the region-based term, and the edge-based term is only used as a stopping force to inhibit the progress of the segmentation boundary across the tumor edges.

The level set methods discussed above evolve the segmentation boundary in an unsupervised fashion. Liu et al. [18] show that, using intensity priors on the white and gray matter tissues, the unsupervised level set method achieves good performance in segmenting the MRI brain images into the two tissue types. But, as discussed previously, in the case of brain tumor segmentation, the lack of shape or intensity priors on the tumors makes it challenging to proceed

in an unsupervised manner. Cobzas et al. [6] presented a MRI tumor segmentation method that incorporates a learned statistical model and atlas-based features into the variational framework. An important aspect of this work was the extraction of multiple features derived from the MRI images using texture information and brain atlases and the integration of the features into the variational segmentation method. A logistic regression density function was used to model the tumor/non-tumor probabilities. In contrast with Ho et al. [28], these probabilities were learned from a set of feature extracted from training data. The evolution was then driven by the difference in these learned tumor/non-tumor probabilities. This method showed good results on MRI images with a wide variety of tumors.

In Table 1, we summarize the above discussed related works on brain tumor segmentation. Here, we observe that most of the existing tumor segmentation methods do not employ priors on the brain shape or appearance. On the other hand, our proposed method incorporates priors on the appearance of normal brain region for a better tumor segmentation performance, as we discuss in the following sections.

Our contribution

In this paper, we propose a supervised 3D variational segmentation method that incorporates additional appearance priors to better disambiguate the tumor from the surrounding deformed brain tissue. We improve upon our earlier work [22] by eliminating the need for manual intervention during the post-processing step (see "Post-processing"). An overview of this segmentation methodology is shown in Fig. 1. Our formulation extends the Chan-Vese region-based segmentation model [5] to incorporate multiple features in a similar way to texture-based approaches [26]. But instead of using an unsupervised approach, we use existing manually labeled data to learn a statistical model and Dirichlet prior from a set of clustered features. Our work is different from the earlier-mentioned standard clustering methods [2,10,34, 37]. These methods attempt a final tumor/non-tumor segmentation through unsupervised classification of the original MRI feature data. Instead, in our work, we use clustering to build a vocabulary (a set of clusters) from training data to parsimoniously represent the salient information in our high-dimensional MRI feature data. The brain tissues and the tumor are then each characterized by one or more of these clusters. The actual segmentation is performed by surface evolution in a supervised manner by the learned inside and outside region statistics in the cluster space. Our method employs a Dirichlet prior during the computation of these voxel probabilities, as opposed to Cobzas et al. [6] where an atlas prior is used only for feature extraction. Also, our method is quite different from the work of Khotanlou et al.

Table 1 Summary of the related brain tumor segmentation approaches

Authors	Description	Training data	Features	Prior	Abnormality
Batista et al. [2]	K-means clustering	No	Texture	No	N/A
Yazdan et al. [37]	K-means clustering	No	DTI	No	N/A
Velthuizen et al. [34]	fuzzy clustering	No	T1, T2, PD	No	MGM, GBM, AST
Fletcher et al. [10]	Knowledge-based fuzzy clustering	Yes	T1, T2, PD	No	Non-enhancing tumors
Vaidyanathan et al. [32]	kNN and fuzzy clustering	Yes	T1, T2, PD	No	GBM
Vinitski et al. [35]	kNN	Yes	T1, T2, PD, T1c	No	N/A
Zhang et al. [38]	One-class SVM	Yes	T1, T1c	No	NMA, CPM, GBM
Kauss et al. [12]	Atlas-based registration with kNN	Yes	SPGR	Yes (Atlas)	MGM, LGG
Prastwa et al. [23]	EM with Atlas priors	No	T1, T1c, T2	Yes (Atlas)	MGM, MLG
Schmidt et al. [27]	SVM	Yes	Texture, atlas, symmetry	No	AST, GBM, OLG
Capelle et al. [3]	Markov random fields	No	T2	No	N/A
Lee et al. [16]	Support vector random fields	Yes	Texture, atlas, symmetry	No	N/A
Lee et al. [17]	Decoupled conditional random fields	Yes	Texture, atlas, symmetry	No	GBM, AST
Corso et al. [7]	Multilevel Bayesian segmentation	No	T1, T1c, T2, FLAIR	Yes	GBM
Archip et al. [1]	Normalized cuts	No	T1, T1c, T2, SPGR	No	MGM, LGG, AST
Droske et al. [9]	Active contour based using level sets	No	N/A	No	LGG, APG
Ho et al. [28]	Active contour based using level sets	No	T1, T1c	No	Enhancing tumors
Xie et al. [36]	Active contour based using level sets	No	SPGR	No	LGG, AST, MGM
Cobzas et al. [6]	Active contour based using level sets	Yes	Texture, symmetry, atlas	No	GBM, AST, MLG, GSM
Proposed approach	Active contour based using level sets	Yes	Texture, symmetry	Yes	GBM, AST, MLG, GSM

The various abbreviations in the above are defined as follows: *DTI* diffusion tensor imaging, *PD* proton density, *SVM* support vector machine, *kNN* *k* nearest neighbor, *MGM* meningioma, *GBM* glioblastoma, *AST* astrocytoma, *NMA* Neuroma, *CPM* craniopharyngioma, *OLG* oligodendroglioma, *SPGR* spoiled gradient recalled, *GSM* gliosarcoma, *LGG* low-grade glioma, *MLG* malignant glioma, *APG* anaplastic glioma, *N/A* information is not available in the paper

Fig. 1 An overview of the proposed tumor segmentation methodology
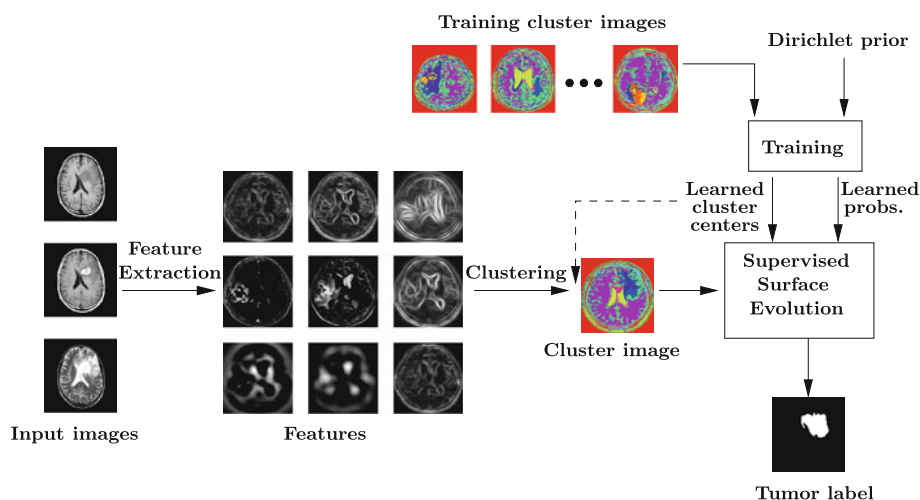
[14], where the active contour framework is employed just to refine the tumor segmentation obtained earlier using symmetry analysis (which is based on the premise that presence of a tumor causes asymmetry between the two hemispheres in the brain). Coincidentally, we also incorporate this symmetry (asymmetry) information into our method, but we compute a symmetry feature for this purpose (see "Feature extraction").

We summarize the main contributions of this paper as follows:

- We extract clusters from a high-dimensional feature set and integrate these clusters into a 3D variational segmentation framework.
- We propose a Dirichlet prior that disambiguates the tumor from the surrounding brain tissue to address the difficult

segmentation cases with tumors that lie close to the major structures in the brain.

- We evaluate our segmentation methodology on real MRI scans from patients with a variety of brain tumors. These experiments show good performance of our method in comparison to the manual segmentation results.

In this paper, we actually consider the segmentation of tumor and also the surrounding *edema* (swelling) region. Henceforth, we use the term "tumor" to refer to the "tumor + edema" region. The rest of this paper is organized as follows, "Theory" describes the basic formulation of the problem and the clustering procedure, "System and implementation details" describes the automatic tumor segmentation system, "Experiments" presents the set of experiments on real MRI data evaluating the performance of our automatic tumor segmentation system. We conclude our discussion in "Conclusions and future work."

Theory

This section describes the general formulation of the 3D variational segmentation problem. Also, we present the clustering procedure that handles multivariate feature data.

Probabilistic formulation of variational segmentation

Given an image volume $I : \Omega \to \Re^+$ defined on a open and bounded domain $\Omega \subset \Re^3$, the binary segmentation task consists of finding a regular surface Γ that splits the domain Ω into two disjoint regions Ω_1, Ω_2. Following [8], an optimal image partitioning $\mathcal{P}(\Omega) = \{\Omega_1, \Omega_2\}$ can be computed by maximizing the *a posteriori* (MAP) partitioning probability $p(\mathcal{P}(\Omega)|I)$ for a given image I. Under a Bayesian assumption:

$$p(\mathcal{P}(\Omega)|I) = p(\{\Omega_1, \Omega_2\}|I) \propto p(I|\{\Omega_1, \Omega_2\}) p(\mathcal{P}(\Omega))$$

(2)

where the first term represents data likelihood and is connected to the region statistics and the second term is a prior (regularization) corresponding to geometric properties of the partition (usually chosen as the length of Γ [8]). Assuming that:

(i) all partitions are equally possible:

$$p(I|\{\Omega_1, \Omega_2\}) = p(I|\Omega_1) \ p(I|\Omega_2)$$

(ii) the voxels within each region are independent:

$$p(I|\Omega_i) = \prod_{\mathbf{x} \in \Omega_i} p_i(\mathbf{x}), \quad i = 1, 2$$

where $p_i(\mathbf{x}) := p(I(\mathbf{x}) = u|\mathbf{x} \in \Omega_i)$, defines the gray value distribution in the region Ω_i.

The above formulation is equivalent to minimizing the following energy [corresponding to the MAP of Eq. (2)]:

$$E(\Omega_1, \Omega_2) = -\int_{\Omega_1} \log \ p_1(\mathbf{x}) \, \mathrm{d}\mathbf{x} - \int_{\Omega_2} \log \ p_2(\mathbf{x}) \, \mathrm{d}\mathbf{x} + \nu|\Gamma|$$

(3)

The optimization of such energy functionals can be efficiently implemented using the level set representation [4,8]. In the level set framework, the segmentation boundary Γ is implicitly represented as a zero-level set or a zero contour of an embedding function $\phi : \Omega \to \Re, \Gamma = \{x \in \Omega | \phi(\mathbf{x}) = 0\}$. The energy functional in (3) can now be expressed as:

$$E(\phi) = \int_{\Omega} \bigg(-H(\phi) \log \ p_1(\mathbf{x}) - (1 - H(\phi)) \log \ p_2(\mathbf{x})$$
$$+ \nu|\nabla H(\phi)| \bigg) \, \mathrm{d}\mathbf{x}$$

(4)

The corresponding Euler–Lagrange evolution equation for ϕ is given by

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \bigg(\log \ p_1 - \log \ p_2 + \nu \, \mathrm{div} \bigg(\frac{\nabla \phi}{|\nabla \phi|} \bigg) \bigg)$$

(5)

where, $H(\phi), \delta(\phi)$ are the Heaviside and Dirac delta functions, respectively.[1]

Remark 1 In the above, the term $(\log \ p_1 - \log \ p_2)$ can be interpreted as the region-based "driving" force. At each pixel, this difference determines the direction of the contour evolution, i.e., if $(\log \ p_1 - \log \ p_2) > 0 (< 0)$ the contour evolves so as to include the pixel in region $\Omega_1(\Omega_2)$.

Region statistics in a multi-dimensional feature space

The probability density functions (PDF) chosen to model p_1, p_2 should: (i) capture the distribution of values in the region and (ii) discriminate between the two regions. Parametric density models (Gaussian, mixture of Gaussians) work well for relatively uniform regions but fail to properly represent more complicated region statistics [26]. We therefore use a Parzen density estimate that can better describe the regions. This method estimates the PDF based on the histograms, employing a Gaussian kernel for smoothing

---

[1] In practice we use the regularized versions of the Heaviside and delta functions, $H_\epsilon(\phi) = \frac{1}{2}\left(1 + \frac{2}{\pi} \tan^{-1}\left(\frac{\phi}{\epsilon}\right)\right)$ and $\delta_\epsilon(\phi) = \frac{1}{\pi} \frac{\epsilon}{\phi^2 + \epsilon^2}$ functions.

$$p_i(s) = \frac{1}{nh} \sum_{k=1}^{n} K\left(\frac{s - s_k}{h}\right), \quad K(s) = \frac{1}{\sqrt{2\pi}} \exp^{-\frac{1}{2}s^2}$$

$$\text{(6)}$$

where s_i, $k = 1 \ldots n$ denote the histogram's gray levels, h is the smoothing parameter.

MRI data usually have more than one sequence (e.g., T1,T2). In addition, to incorporate the texture information in the MRI images, we extract a set of features that better capture the image local scale and frequency (similar to Rousson et al. [26]) using Gabor-type filters. The particular features used are shortly presented in "Feature extraction." We therefore need to generalize the above formulation to vector valued data.

Let $\mathcal{I} = \{I_1, I_2, \ldots, I_m\}$ be the set of feature images. If we assume the features to be independent of each other [5,8,26], the total likelihood for the region Ω_i $(i = 1, 2)$ is

$$p(\mathcal{I}|\Omega_i) = \prod_{j=1}^{m} p\left(I_j | \Omega_i\right) = \prod_{j=1}^{m} \prod_{\mathbf{x} \in \Omega_i} p_{i,j}(\mathbf{x}) \quad \text{(7)}$$

The independence assumption might not be valid in the case of significant correlation among features. In such a situation, one can consider the set features as multivariate data $\mathbf{I} : \Omega \to \Re^m$, where each pixel location corresponds to a m dimensional vector $\mathbf{I}(\mathbf{x}) = [I_1(\mathbf{x}) \ I_2(\mathbf{x}) \cdots I_m(\mathbf{x})]$ [25]. But estimating high-dimensional densities, in particular the nonparametric density is computationally intensive. We propose a clustering approach to handle the multivariate image data. The idea of clustering features was previously used by Malik et al. [20] to identify texons (texture patterns) in a feature space in the context of discrete segmentation (Normalized cuts).

We use the K-means algorithm [19] with a Euclidean distance measure in the space of the features to obtain K cluster centers $\mathbf{c}_k \in \Re^m$, $k = 1, 2, \ldots, K$. We define a "cluster image," $\hat{I} : \Omega \to \{1, 2, \ldots, K\}$ corresponding to a given vector image $\mathbf{I}$ as

$$\hat{I}(\mathbf{x}) = \underset{k}{\operatorname{argmin}} \parallel \mathbf{I}(\mathbf{x}) - \mathbf{c}_k \parallel^2 \quad \forall \mathbf{x} \in \Omega \quad \text{(8)}$$

i.e., we assign each pixel in the vector image to its nearest cluster center. A set of feature images is thus replaced by a single "cluster image" (see Fig. 2) in the active contour segmentation model. The evolution equation is then given by the general active contour evolution in (5), with p_1, p_2 now defined in the cluster space. We mention two important differences between our proposed active contour evolution (in the cluster space) and the general active contour evolution in (5):

- As the range of cluster values ($k \in \{1, 2, \ldots, K\}$) is finite, *discrete* PDFs are used to model p_1, p_2. Specifically, we
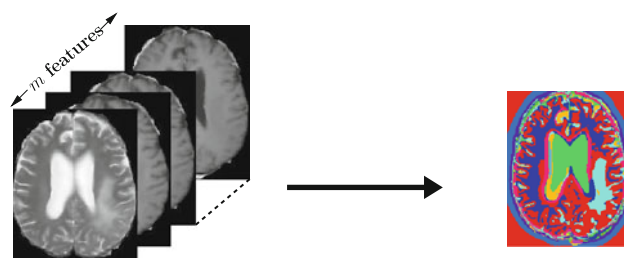

Fig. 2 Features mapped to a cluster image via K-means clustering

choose the multinomial density function:

$$p\left(\hat{I}(\mathbf{x}) = k | \mathbf{x} \in \Omega_i, \boldsymbol{\theta}^i\right) = p_i\left(\hat{I}(\mathbf{x}) = k | \boldsymbol{\theta}\right) = \theta_k^i$$

$$\text{where} \quad \boldsymbol{\theta}^i = \left(\theta_1^i, \theta_2^i, \ldots, \theta_K^i\right) \quad \text{and} \quad \sum_{k=1}^{K} \theta_k^i = 1$$

$$i \in \{1, 2\} \quad \text{(9)}$$

i.e., parameter θ_k^i represents the probability of a pixel (in a region Ω_i) belonging to the cluster k.

- The region statistics in the "cluster image," which are basically the cluster value probabilities (parameterized by $\boldsymbol{\theta}^i$s) are <u>not</u> updated after every iteration of the contour evolution, instead they are pre-computed (along with the cluster centers $\mathbf{c}_k$) from a set of training images.

Separating tumor and normal brain using a Dirichlet prior

Most variational segmentation techniques are used in an unsupervised setting where the region statistics are refined as the curve evolves [4,26]. This might be quite effective if the region statistics are distinct. But, as mentioned earlier, one of the main problems in brain tumor segmentation is that the appearance of tumor and surrounding tissue is not always clearly separated (not even in feature space). As an example, see Fig. 6. We therefore have to use additional prior information to help the segmentation. The tumor does not have a particular shape prior. In addition, the surrounding tissues (like the ventricles) can be deformed and therefore do not preserve a shape prior. We chose to use a prior on the appearance that better disambiguates the two regions.

Most segmentation errors in our automatic system are caused by the vicinity of tumor and the normal structures in the brain (like ventricles, eyes, etc.) when part of the normal brain region is incorrectly segmented as tumor (see Fig. 6). We designed a prior that penalizes the clusters predominant in the normal brain region from having a high probability in the tumor. Hence, we assumed a Dirichlet prior for (p_1, p_2) skewed in such a fashion that those prominent clusters in the normal brain region have a very low prior probability, and the rest of the clusters have uniform probability. We used images

with manually segmented tumor to identify the normal brain and tumor regions.

Given $\mathcal{D} = \{\hat{I}^{(1)}, \hat{I}^{(2)}, \ldots, \hat{I}^{(N)}\}$, a set of cluster images[2] with labeled tumor and normal brain regions, we need to estimate the cluster value probability distribution in the region Ω_i, denoted as $p_i(\hat{I}(\mathbf{x})|\mathcal{D})$. We know from (9) that p_i is modeled as a multinomial with parameters $\boldsymbol{\theta}^i = (\theta_1^i, \theta_2^i, \ldots, \theta_K^i)$. In a Bayesian learning approach, we assume $\boldsymbol{\theta}^i$ to have a *prior* distribution $p(\boldsymbol{\theta}^i)$ and our goal is to compute the *posterior* probability $p(\boldsymbol{\theta}^i|\mathcal{D})$ given the observations. We choose a *Dirichlet* distribution to model such parameter priors:

$$p(\boldsymbol{\theta}^i) = \mathrm{Dir}\left(\alpha_1^i, \alpha_2^i, \ldots, \alpha_K^i\right) \sim \prod_{k=1}^{K}\left(\theta_k^i\right)^{\alpha_k-1} \quad (10)$$

where $\{\alpha_1^i, \alpha_2^i, \ldots \alpha_K^i\}$ are called the hyperparameters. Our prior knowledge about θ_k^i is represented by the corresponding hyperparameter α_k^i. Now, we can compute the posterior based on the property of Dirichlet distribution [15] as:

$$p(\boldsymbol{\theta}^i|\mathcal{D}) = \mathrm{Dir}\left(\alpha_1^i + M_1^i, \alpha_2^i + M_2^i, \ldots, \alpha_K^i + M_K^i\right) \quad (11)$$

where M_k^i is the number of pixels with a cluster value k that occur in the region Ω_i among the training images. We can estimate the cluster value distribution in the regions as:

$$p_i\left(\hat{I}(\mathbf{x}) = k|\mathcal{D}\right) = \int p_i\left(\hat{I}(\mathbf{x}) = k|\mathcal{D}, \boldsymbol{\theta}^i\right) \; p\left(\boldsymbol{\theta}^i|\mathcal{D}\right) \mathrm{d}\boldsymbol{\theta}^i$$

$$= \int \theta_k^i \; p\left(\boldsymbol{\theta}^i|\mathcal{D}\right) \; \mathrm{d}\boldsymbol{\theta}^i = \mathbf{E}_{p\left(\boldsymbol{\theta}^i|\mathcal{D}\right)}\left[\theta_k^i\right] \quad (12)$$

Using the expectation of a Dirichlet distribution, we get:

$$p_i\left(\hat{I}(\mathbf{x}) = k|\mathcal{D}\right) = \frac{\alpha_k^i + M_k^i}{\sum_l \alpha_l^i + \sum_l M_l^i} \quad (13)$$

From the training data we have prior knowledge that a certain cluster k is less likely to occur in a region Ω_i. By choosing a relatively low value for α_k^i, we can suppress the probability of this cluster k in the region Ω_i. This is illustrated in Fig. 3, where the cluster value dominant in the normal brain region is assigned a low probability in the tumor region, after estimation with a biased prior (here for clusters 7, 10 we chose, $\alpha_7^1 = \alpha_{10}^1 = 10^{-4}$ and for the other clusters $k \notin \{7, 10\}$, $\alpha_k^1 = 10^6$). In Fig. 3, the "Before" plot represents the estimated probabilities without the use of a Dirichlet prior and the "After" plot represents the estimated probabilities with the use of Dirichlet prior. Here, we observe that the normal brain clusters 7, 10 are assigned a low prior probability (according to the Dirichlet prior). Thus, in the "After" plot these normal brain clusters are shown to "suppressed," i.e., they are less

---

[2] We consider each pixel in the set of cluster images to be an independent and identically distributed (i.i.d.) sample.
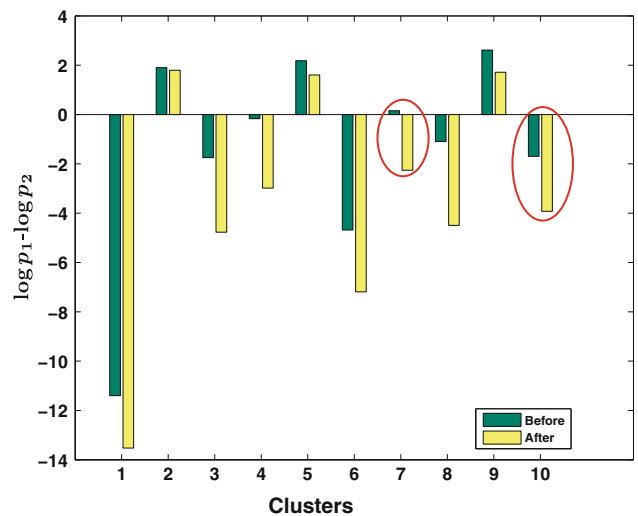

Fig. 3 Suppression of normal brain clusters: the clusters 7, 10, which are the two most prominent clusters in the normal brain region histogram are suppressed, notice that for the clusters 7, 10 it is ensured that $(\log p1 - \log p2) < 0$ after suppression

Training:
    (1) pre-processing and feature extraction
    (2) clustering features
    (3) compute normal brain prior and
        PDF for 'tumor' p_1 and 'brain' p_2 (Section 2.2,2.3)
Segmentation:
    Init: (1) pre-processing and feature extraction
        (2) assign clusters and init level set
    Evolve: (3) evolve level set until converg.(Equation 5)
    Finish: (4) remove pieces of the skull and parts of the CSF

Fig. 4 Overview of the segmentation system

likely to be in the tumor region as shown by the negative value of $\log p1 - \log p2$.

System and implementation details

The raw MRI data for each patient consist of the T1, T1 with contrast (T1c), and T2 image volumes. For both training and testing, we first pre-process the images. Next we extract the feature set. We use Gabor-type texture features and a left–right symmetry feature along with the original image modalities. We cluster these extracted features and learn the PDFs on the brain/tumor regions. In this section, we present details on our segmentation system.

Segmentation system

The system is briefly summarized in Figs. 1 and 4. The training phase follows with clustering of the extracted features and calculation of brain/tumor PDFs as described in Sects. "Region statistics in a multi-dimensional feature

space," and "Separating tumor and normal brain using a Dirichlet prior." For segmenting a new image, we first assign clusters and initialize the level set with a circle in the middle of the image. The curve is then evolved (Eq. 5) until it reaches a stable position. In a post-processing step, after the evolution is finished, we automatically remove the pieces of skull and parts of CSF that were mislabeled as tumor. The following briefly describe the steps involved in the pre-processing, post-processing and feature extraction of the MRI images.

Data pre-processing

The various pre-processing steps are briefly described below:

1. **Noise reduction** The images are smoothed using an edge preserving non-linear filter [31].
2. **Intensity standardization** The inter-slice intensity variation is corrected using weighted least squares. The intensity inhomogeneity across the volume is reduced by the N3 method [30].
3. **Symmetry alignment** In this step, we correct for the misalignment of the MRI images about the y-z plane, i.e., when the plane of symmetry in the MRI image does not coincide with the y-z plane. This can result in errors while computing the symmetry feature (see "Feature extraction"). We achieve symmetry alignment by applying a rigid transformation to the image. The transformation parameters are computed as "half" of the transformation parameters obtained from the affine registration of the image with its mirror reflection about the y-z plane.

Feature extraction

The original data comprised of different image modalities are mapped to a high-dimensional feature space where each image voxel corresponds to a m-dimensional vector. The feature volumes are generated by stacking the features extracted on the 2D image slices. We consider three sets of features. Figure 5 shows the extracted feature images.

Image modality features: The T2 images were directly included into the feature set. The T1c images (T1, with gadolinium contrast agent) highlight the abnormal "enhancement" regions, indicating the presence of a tumor. This motivated the use of image difference between the T1c and T1 modalities (T1–T1c) as a feature.

Gabor features: The tumor generally has quite different texture characteristics from the normal brain tissue. We extract texture features using the MR8 (Gabor-type) [33] filter bank responses. They essentially contain an edge and a bar filter, each at 6 orientations and 3 scales (coarse, medium, fine). Only the maximum response over the 6 orientations
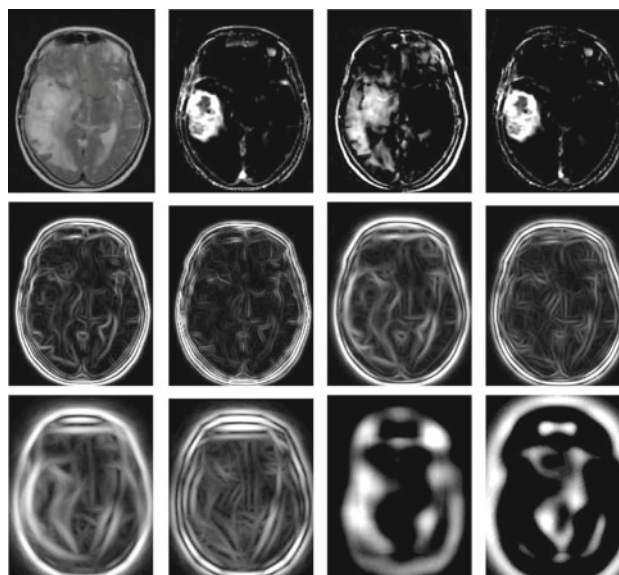

Fig. 5 Feature images: (*top row*, l–r) T2, T1-T1c, symmetry (T2), symmetry (T1-T1c); (*middle row*, l–r) fine- and medium-scale Gabor features; (*bottom row*, l–r) coarse-scale Gabor features, Gaussian, Laplacian of Gaussian

is stored resulting in 6 features. The maximum response is just the maximum intensity value at each pixel obtained over the 6 orientations. The Gaussian and Laplacian of Gaussian responses are also incorporated, resulting in a total of 8 filter responses. We compute the Gabor features using T2 images only.

Symmetry feature: In general, the two hemispheres of the normal brain are symmetric about the central line of symmetry. The occurrence of tumors (most of which are observed to be localized in one of the brain hemispheres) results in a left–right asymmetry. We compute a feature characterizing the symmetry in the brain by taking corresponding pixel intensity differences about the line of symmetry. We compute two symmetry features using the T2 and T1–T1c images, respectively.

In our final experiments, we only considered the fine-scale Gabor features (see Fig. 5, middle row, col 1–2) along with the image modality and symmetry features (see Fig. 5, top row, col 3–4), the Gaussian and the Laplacian of Gaussian (see Fig. 5, bottom row, col 3–4).

Post-processing

In an automatic post-processing step after the evolution is finished, we remove pieces of the skull and parts of CSF that were wrongly labeled as being part of the tumor region. For this purpose, we perform a "crude" segmentation of the skull and CSF regions in the MRI images. Then, we refine the tumor label using these skull and CSF segmentation masks. Further, as most of the tumor exists as a single piece in the
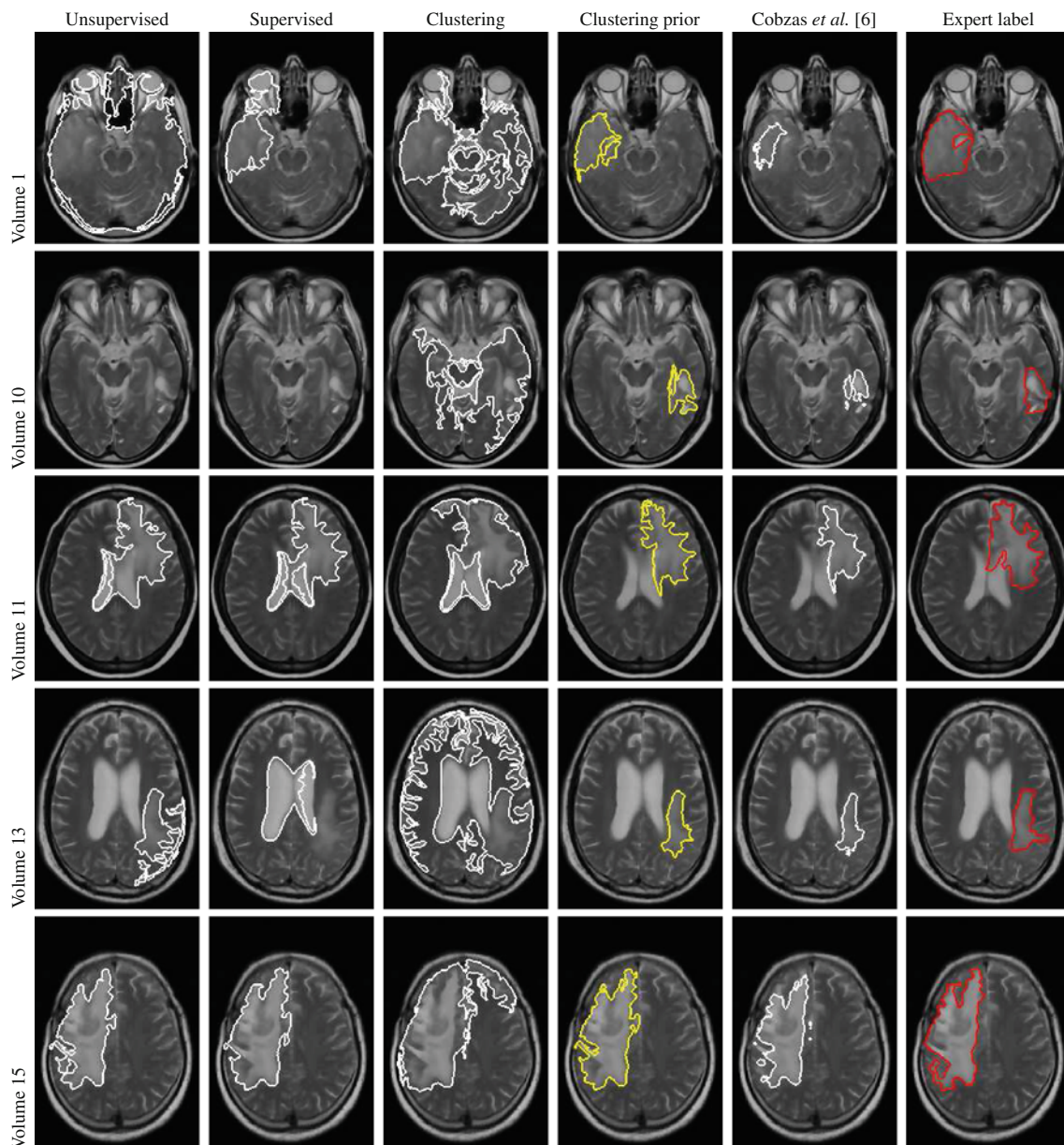
Fig. 6 Comparative results of different segmentation methods (rows represent different MRI volumes). col 1 *white*—unsupervised method with feature histograms (automated label); col 2 *white*—supervised method with feature histograms (automated label); col 3 *white*—clustering method (automated label); col 4 *yellow*—clustering method + ventricle prior (automated label); col 5 *white*—method proposed by Cobzas et al. [6] (automated label); col 6 *red*—(manual label)

brain volume and also on each of the corresponding 2D slices, we employ a heuristic to pick the largest piece to obtain the final tumor segmentation label. The post-processing procedure is described below:

(1) **Skull removal**: For a given MRI volume to be segmented, we performed unsupervised clustering on the corresponding extracted feature data. We then obtained the skull region automatically as belonging to two particular clusters. The MRI features we considered for skull removal were: the image modality features and the symmetry features, the Laplacian of Gaussian feature. In addition, we incorporated a "distance" feature that gives the distance of each point in the image to the boundary of the whole brain region. We computed this feature on every 2D slice of the MRI volume. We used the K-means clustering method and chose $K = 6$ clusters. We found that 2 of these clusters adequately described the skull region. These 2 clusters were automatically chosen as they correspond to the top two cluster centers that had the highest values for the "distance" feature channel.

(2) **CSF removal**: To segment the CSF, we only considered the T1 MRI image. We performed contrast enhancement on the T1 image using the standard CLAHE method [39]. The CSF was then segmented using simple thresholding. The threshold value was fixed at the beginning of our experiments and it was chosen to be the same for all the MRI images in our data set.

(3) **Choose the largest piece**: We improved this automated segmentation label further by choosing the largest 2D connected component on each slice and then ensuring 3D consistency by picking the largest 3D component.

(4) **Hole filling**: In addition to the above steps, we filled any "holes" (small regions inside the tumor, labeled as normal by our algorithm) in the tumor label.

We emphasize that the above described post-processing steps are completely automatic and require no user input, except for setting the initial threshold for CSF segmentation.

We also considered the idea of removing skull and CSF as a pre-processing step. However, we hypothesize that the skull and CSF removal as a post-processing step results in a better performance for our method. This is because, if we perform the skull and CSF removal operations as a pre-processing step, there is a chance that some of the misclassified CSF/skull voxels (false positives) could fall into the tumor area and would therefore not be considered in the training data. This would introduce a bias in the training data influencing the learning of cluster centers, Dirichlet prior and tumor/brain PDFs leading to a poor performance of our proposed method. We verified our hypothesis by an experiment where we compared the two approaches (pre- and post-skull and CSF removal) on a subset of 6 MRI volumes and found that our proposed method performed significantly better [$\approx$ 20% improvement in Jaccard scores (defined in Table 3)] when we employed the skull and CSF removal as a post-processing step rather than a pre-processing step. Hence, we chose skull and CSF removal as a post-processing step.

Experiments

In this section, we evaluate the performance of our proposed automatic tumor segmentation system on real MRI data through two experiments. The first experiment is designed to tune the clustering component of the system, i.e., we determine the optimal number of clusters to be used in the K-means clustering step of our proposed segmentation method. In the second experiment, we compare our proposed approach with other variational segmentation methods.

Our experimental data set (obtained from the Cross Cancer Institute, Alberta) consists of 15 MRI volumes of patients with brain tumors. The MRI volumes were available in three sequences: T1-weighted (T1), T1 with gad-

Table 2 Tumor type information associated with the MRI volumes in our data set

Volume no.	Tumor type
1	Astrocytoma
2	Astrocytoma
3	Glioblastoma
4	Glioblastoma
5	Glioblastoma
6	Glioblastoma
7	Glioblastoma
8	Glioblastoma
9	Glioblastoma
10	Glioblastoma
11	Glioblastoma
12	Gliosarcoma
13	Glioma, malignant
14	Glioblastoma
15	Glioblastoma

Table 3 The various metrics used to evaluate the performance our proposed segmentation algorithm

Jaccard score	$\mathrm{Jacc}(V_A, V_E) = \frac{tp}{tp+fp+fn}$
Precision	$\mathrm{Prec}(V_A, V_E) = \frac{tp}{tp+fp}$
Recall	$\mathrm{Rec}(V_A, V_E) = \frac{tp}{tp+fn}$
Hausdorff distance	$\mathrm{Haus}(S_A, S_E) = \max\{\sup_{a \in S_A} \inf_{e \in S_E} d(a, e), \sup_{e \in S_E} \inf_{a \in S_A} d(a, e)\}$

In the above, V_A, V_E denote the automatic and expert segmentation volumes, respectively, S_A, S_E denote the automatic and expert segmentation 3D surfaces, respectively, tp denotes the number of true positive voxels, fp denotes the number of false positive voxels, fn denotes the number of false negative voxels, $d(a, e)$ denotes the Euclidean distance between any two points $a \in S_A$ and $e \in S_e$

olinium contrast agent (T1c) and T2-weighted (T2). Each image sequence contains a series of 21 consecutive 2D image slices of the brain taken with a 6.5 mm slice-interval (i.e., the voxel size was 1 mm $\times$ 1 mm $\times$ 6.5 mm). In Table 2, we list the tumor type associated with each of these MRI volumes. Our data set contains of tumors with varied size, location and intensity characteristics, for instance see Fig. 6: small tumors (Volume 10), tumors growing very close to the ventricles (Volumes 3, 11, 13), non-homogeneous tumors (Volume 3), tumors growing close to the eyes (Volume 1) and very large tumors (Volume 15). The tumor region was manually segmented in each MRI volume by expert radiologists. These expert labels were used for training and validation (in a leave-one-out manner).

Number of clusters

The goal of this experiment is to determine the number of clusters (K) required to sufficiently describe the high-dimensional MRI feature data. Clearly, a proper choice of the parameter K is necessary for the optimal performance of our proposed segmentation method. The parameter K should be chosen such that the tumor and the brain regions are sufficiently characterized and well separated in the cluster space. The major classes in the normal brain are the gray matter (GM), white matter (WM), cerebrospinal fluid (CSF) and the bone (Skull). Therefore, ideally we should be able to properly represent the normal brain region using only 4 clusters. But, due to the significant variability within each of the above classes, often more than one cluster per class is needed. Further, the presence of tumors and the surrounding edema adds to this existing variability in the brain.

We employed a brute force approach to determine the value of K that results in an optimal performance of our automatic segmentation method. For this purpose, we randomly chose a small subset of 6 MRI volumes from the original 15 MRI volumes. We ran our proposed automatic segmentation method on these 6 MRI volumes using different values of K, i.e., $K = 4, 5, \ldots, 20$. Each of the segmentation runs was performed using the leave-one-out cross-validation framework. In each run, we compared the tumor segmentation label obtained using our automatic segmentation method with the expert label and recorded the average (over the 6 MRI volumes) Jaccard score (defined in Table 3). We obtained the highest average Jaccard score for $K = 10$. Therefore, we choose $K = 10$ as the optimal number of clusters.

Comparison with other segmentation methods

In this experiment, we want to quantitatively determine the improvement in the tumor segmentation results achieved by our proposed method through the use of training data, clustered features and prior knowledge, respectively. Further, we also want to evaluate the performance of our proposed method in the context of existing variational tumor segmentation methods. Hence, we compare the segmentation results obtained using the following 5 variational segmentation methods:

1. traditional unsupervised segmentation with Parzen histograms on the feature set images [8,26].
2. supervised segmentation on the feature images, where Parzen histograms were computed based on the training data.
3. supervised segmentation on the clustered image.
4. our proposed segmentation method that uses appearance priors to disambiguate the tumor from the normal brain tissues.

5. variational tumor segmentation method proposed by Cobzas et al. [6].

The first method (popularly used for texture segmentation [8,26]) assumes that the set of features are independent and estimates Parzen histograms in both the tumor and normal regions at each evolution step. The second method takes a supervised approach and computes the Parzen histograms based on the training data and performs supervised segmentation on the set of assumed independent feature images. The third method discards the assumption of feature independence and instead uses our proposed clustering-based active contour model. However, this method does not incorporate prior knowledge about the normal brain tissues during the estimation of cluster value probabilities in the tumor region. The fourth method is our proposed segmentation method that uses learned cluster value probabilities along with appearance priors to better discriminate the tumor from the normal brain tissues. The fifth method proposed by Cobzas et al. [6] employs a multi-dimensional atlas-based feature set in a 3D variational framework. For each of the 1–4 segmentation methods, the regularization parameter ν [see (5)] was chosen to be the same and was fixed at the beginning of the experiments. Further, the automated tumor labels from each of the 1–4 segmentation methods were refined using identical post-processing steps (as discussed in "Post-processing"). For the method 5 (by Cobzas et al. [6]), we obtained an implementation of the method from the authors.

We performed segmentation of the tumor region in all the 15 MRI volumes in our data set using each of the above mentioned 5 automatic segmentation methods, in a leave-one-out cross-validation fashion. To compare the segmentation results obtained, we considered 4 different metrics, namely the Jaccard score, Precision, Recall and the Hausdorff distance (defined in Table 3). The segmentation results are tabulated in Table 4. We see that our proposed segmentation method achieves the highest mean Jaccard score and the lowest mean Hausdorff distance. Further, our method obtains a significantly higher precision score (at least 22% higher) on average compared to the other segmentation methods, at a reasonable recall rate.

In Fig. 6, we show 2D slices illustrative of the performance of the five segmentation methods (cols 1–5) for some of the MRI volumes in our experimental data set. These MRI volumes shown in Fig. 6 were chosen such that they are representative of the variety of tumors present in our MRI data set. Further, in Fig. 7 we show a 3D segmentation label corresponding to volume 15 for illustration. In the case of volume 13, where the tumor lies close to the ventricles, our method clearly discriminates between the tumor region and the ventricle region, whereas the segmentation methods 1–3 erroneously label the ventricles and/or the CSF surrounding the tumor as tumor. Subsequently in this case, our

Table 4 Comparison of the Jaccard score (%), Precision (%),Recall (%) and the Hausdorff distance (mm) obtained using different segmentation methods

Scan no.	Unsupervised				Supervised				Clustered feat.				Clustered feat. with prior				Cobzas et al. [6]			
	Jacc	Prec	Rec	Haus	Jacc	Prec	Rec	Haus	Jacc	Prec	Rec	Haus	Jacc	Prec	Rec	Haus	Jacc	Prec	Rec	Haus
1	4	4	99	95	45	52	76	79	7	7	97	95	**53**	**82**	**60**	**20**	33	34	92	19
2	50	59	76	44	46	48	88	51	31	32	94	73	**49**	**58**	**75**	**44**	49	65	66	43
3	8	9	87	94	73	80	88	42	57	65	82	74	**62**	**89**	**67**	**15**	49	50	94	19
4	2	2	99	121	11	11	71	119	4	4	95	119	**51**	**84**	**57**	**32**	45	52	77	30
5	48	65	64	47	56	69	74	48	63	78	76	26	**60**	**90**	**64**	**20**	53	56	90	16
6	72	94	75	13	68	79	83	47	52	55	90	76	**73**	**92**	**78**	**13**	65	66	97	13
7	59	89	63	16	68	81	81	35	42	43	95	91	**64**	**88**	**71**	**15**	49	51	93	18
8	53	87	58	51	54	83	61	40	56	65	80	33	**55**	**90**	**58**	**56**	33	33	97	58
9	46	51	83	71	51	65	70	48	20	21	82	85	**47**	**66**	**62**	**27**	28	29	87	24
10	0	0	0	130	0	0	0	130	1	1	94	133	**50**	**68**	**65**	**24**	25	40	40	44
11	65	71	88	43	67	74	87	43	37	37	96	84	**66**	**81**	**78**	**17**	60	65	89	13
12	49	98	49	19	67	96	69	19	48	60	71	67	**61**	**98**	**62**	**19**	54	54	99	19
13	34	34	93	96	30	35	68	68	8	8	81	102	**63**	**82**	**73**	**17**	68	76	86	14
14	42	50	73	54	41	50	68	80	1	2	6	96	**51**	**71**	**63**	**36**	13	14	79	54
15	76	88	84	17	71	76	90	43	24	24	98	100	**73**	**87**	**81**	**14**	64	67	93	15
Mean	40	53	72	60	49	59	71	59	30	33	82	83	**58**	**81**	**67**	**24**	45	50	85	26
Std.	25	35	24	38	22	27	21	30	22	26	22	28	**8**	**11**	**7**	**12**	16	17	15	15

It can be seen that our proposed method (Clustered features with prior) has the best performance, with the highest mean Jaccard score and the lowest mean Hausdorff distance
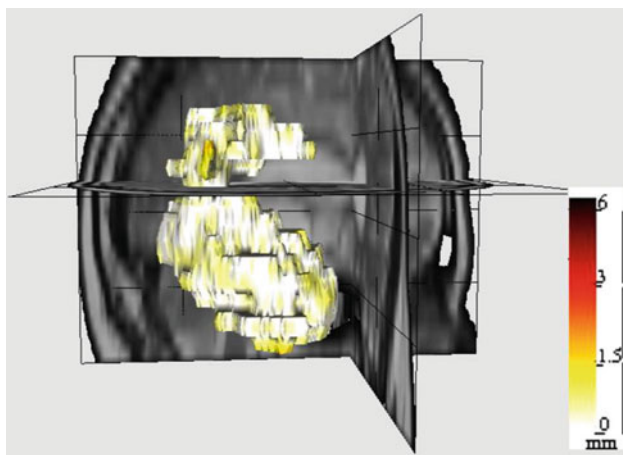

Fig. 7 Volume 15, 3D tumor surface with color-coded distance error. Here, the actual (expert labeled) 3D tumor surface is shown. The color denotes the deviation of the automated segmentation surface generated by our proposed method from the actual tumor surface. The deviation is computed as the distance between the two surfaces

proposed method achieves a considerably higher Jaccard score and a considerably lower Hausdorff distance compared to the segmentation methods 1–3. Further, the performance of our method in this case is similar to the performance of the method by Cobzas et al. [6]. Again, in the case of volume 11 we have a situation where the tumor is near the ventricles. Even here, our method is successful in disambiguating the ventricles from the tumor region, whereas the methods 1–3 confuse part of the ventricles as tumor. But, in this case the Jaccard scores obtained using our method are similar to the Jaccard scores from the unsupervised and the supervised methods. However, the better performance of our method is obvious from the significantly lower Hausdorff distance between the expert tumor label and the automated tumor label generated by our method. We also observe that in this case our method obtains a slightly better Jaccard score than the method by Cobzas et al. [6]. In the case of MRI volume 1, where the tumor is present near the eyes, again our method successfully identifies the eyes as part of the normal brain and achieves a good tumor segmentation label. The other segmentation methods (including the method by Cobzas et al. [6]) show poor performance and the supervised method (see Fig. 6, volume 1: col 2) particularly fails to distinguish the eyes from the tumor. We have contrasting situations of a small tumor and a large tumor in the cases of MRI volume 10 and MRI volume 15, respectively. In both these cases, our method shows a good segmentation performance. However, the other methods (including the method by Cobzas et al. [6]) perform poorly in the case of the small tumor in MRI volume 10. In fact, the unsupervised and supervised methods (see Fig. 6, volume 10: cols 1, 2) completely fail in identifying the tumor region (They obtain a Jaccard score of 0). This is because, in this case the unsupervised and supervised methods mislabeled large regions of the normal brain

as tumor, and consequently in the post-processing step as we pick the largest piece as tumor, the falsely labeled normal brain regions were selected instead of the relatively smaller "actual" tumor region.

Conclusions and future work

We presented an automatic method for the segmentation of tumor + edema from brain MRI images. Existing region-based variational segmentation methods are not suited for tumor segmentation as they are not discriminative enough when the appearance of tumor and normal tissue overlap. We proposed a novel supervised 3D variational segmentation method. This method uses priors on the brain/tumor appearance calculated on a set of clustered features extracted from the MRI images and was able to better disambiguate the tumor from the surrounding tissue.

We have identified a scope for improvement in the MRI feature extraction step. In our current method, we use the standard texture features (Gabor-type). These features were originally designed for natural images with rich texture, hence they have limited success with the gray-scale MRI images with relatively less texture information. There exist various other image features like structure tensor (gradient-based), Leung-Malik and Schmid Filter Bank (Gabor-based). Of particular interest are the features extracted based on the "field of experts" (FoE) model [24]. In this approach, we train the FoE model on a MRI image database and thus adaptively learn a set of image filters, instead of using the generic set of texture filters.

Conflict of interest   None.

References

1. Archip N, Jolesz F, Warfield S (2007) A validation framework for brain tumor segmentation. Acad Radiol 14(10):1242–1251
2. Batista J, Kitney R (1995) Extraction of tumors from MR images of the brain by texture and clustering. In: Proceedings of the 8th international conference on image analysis and processing (ICIAP '95). Springer, London, pp 235–240
3. Capelle AS, Alata O, Fernandez-Maloigne C, Ferrié JC (2000) Unsupervised segmentation for automatic detection of brain tumors in MRI. In: IEEE international conference on image processing, ICIP 2000, Vancouver, Canada
4. Chan T, Vese L (2001) Active contours without edges. IEEE Trans Image Process 10(2):266–277
5. Chan T, Sandberg B, Vese L (2000) Active contours without edges for vector-valued images. J Vis Commun Image Represent 11(2):130–141
6. Cobzas D, Birkbeck N, Schmidt M, Jagersand M, Murtha A (2007) 3d variational brain tumor segmentation using a high dimensional feature set. In: IEEE 11th international conference on computer vision, pp 1–8
7. Corso J, Sharon E, Dube S, El-Saden S, Sinha U, Yuille A (2008) Efficient multilevel brain tumor segmentation with inte-

8. grated Bayesian model classification. IEEE Trans Med Imaging 27(5):629–640
8. Cremers D, Rousson M, Deriche R (2007) A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. Int J Comput Vis 72(2):195–215
9. Droske M, Bernhard M, Martin R, Carlo S (2001) An adaptive level set method for medical image segmentation. In: IPMI '01: Proceedings of the 17th international conference on information processing in medical imaging, Springer, London, pp 416–422
10. Fletcher-Heath L, Hall L, Goldgof D, Murtagh F (2001) Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. Artif Intell Med 21(1-3):43–63
11. Gibbs P, Buckley D, Blackb S, Horsman A (1996) Tumour volume determination from MR images by morphological segmentation. Phys Med Biol 41:2437–2446
12. Kaus M, Warfield S, Nabavi A, Black P, Jolesz F, Kikinis R (1999) Adaptive template moderated brain tumor segmentation in MRI. Bildverarbeitung fur die Medizin. Springer, Berlin, pp 102–106
13. Kaus M, Warfield S, Nabavi A, Simon K, Chatzidakis E, Black P, Jolesz F, Kikinis R (1999) Segmentation of meningiomas and low grade gliomas in MRI. In: MICCAI '99: Proceedings of the second international conference on medical image computing and computer-assisted intervention, Springer, London, pp 1–10
14. Khotanlou H, Atif J, Colliot O, Bloch I (2006) 3D brain tumor segmentation using fuzzy classification and deformable models. Fuzzy logic and applications. Springer, New York, pp 312–318
15. Koller D, Friedman N (2005) Structured probabilistic models, (unpublished manuscript)
16. Lee C, Greiner R, Schmidt M (2005) Support vector random fields for spatial classification. Lect Notes Comput Sci 3721:121
17. Lee C, Greiner R, Zaiane O (2006) Efficient spatial classification using decoupled conditional random fields. Lect Notes Comput Sci 4213:272
18. Liu J, Chelberg D, Smith C, Chebrolu H (2007) Distribution-based level set segmentation for medical images. In: 18th British machine vision conference (BMVC'07). Warwick, UK, 10–13 Sept, 2007
19. Macqueen JB (1967) Some methods of classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematics, statistics and probability, pp 281–297
20. Malik J, Belongie S, Leung T, Shi J (2001) Contour and texture analysis for image segmentation. IJCV 43(1):29–44
21. Mazzara G, Velthuizen R, Pearlman J, Greenberg H, Wagner H (2004) Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. Int J Radiat Oncol Biol Phys 59(1):300–312
22. Popuri K, Cobzas D, Jagersand M, Shah S, Murtha A (2009) 3D variational brain tumor segmentation on a clustered feature set. In: Society of photo-optical instrumentation Engineers (SPIE) conference series
23. Prastawa M, Bullitt E, Moon N, Leemput K, Gerig G (2003) Automatic brain tumor segmentation by subject specific modification of atlas priors. Acad Radiol 10(12):1341–1348
24. Roth S, Black M (2005) Fields of experts: a framework for learning image priors. In: IEEE computer society conference on computer vision and pattern recognition, IEEE Computer Society, 1999, vol 2, p 860
25. Rousson M, Deriche R (2002) A variational framework for active and adaptative segmentation of vector valued images. In: IEEE workshop on motion and video comp, pp 56–61
26. Rousson M, Brox T, Deriche R (2003) Active unsupervised texture segmentation on a diffusion based feature space. In: Computer vision and pattern recognition, 2003. Proceedings. 2003 IEEE computer society conference on, vol 2, pp 699–704
27. Schmidt M, Levner I, Greiner R, Murtha A, Bistritz A (2005) Segmenting brain tumors using alignment-based features. In: Pro-

ceedings fourth international conference on machine learning and applications, 2005. pp 215–220

28. Sean H, Elizabeth B, Guido G (2002) Level set evolution with region competition: automatic 3-d segmentation of brain tumors. In: Proceedings of the 16th international conference on pattern recognition. IEEE Computer Society, pp 532–535

29. Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22:888–905

30. Sled JG, Zijdenbos AP, Evans AC (1998) A nonparametric method for automatic correction of intensity nonuniformity in mri data. IEEE Trans Med Imaging 17(1):87–97

31. Smith S, Brady J (1997) Susan a new approach to low level image processing. Int J Comput Vision 23(1):45–78

32. Vaidyanathan M, Clarke L, Velthuizen R, Phuphanich S, Bensaid A, Hall L, Bezdek J, Greenberg H, Trotti A, Silbiger M (1995) Comparison of supervised MRI segmentation methods for tumor volume determination during therapy. Magn Reson Imaging 13(5):719–728

33. Varma M, Zisserman A (2005) A statistical approach to texture classification from single images. Int J Comput Vis V62(1):61–81

34. Velthuizen R (1995) Validity guided clustering for brain tumor segmentation [treatment planning]. In: IEEE 17th annual conference engineering in medicine and biology society, 1995, vol 1, pp 413–414

35. Vinitski S, Gonzalez C, Knobler R, Andrews D, Iwanaga T, Curtis M (1999) Fast tissue segmentation based on a 4D feature map in characterization of intracranial lesions. J Magn Reson Imaging 9(6):768–776

36. Xie K, Yang J, Zhang Z, Zhu Y (2005) Semi-automated brain tumor and edema segmentation using MRI. Eur J Radiol 56(1): 12–19

37. Yazdan-Shahmorad A, Jahanian H, Patel S, Soltanian-Zadeh H (2007) Automatic brain tumor segmentation using tissue diffisivity characteristics. In: ISBI 2007 4th IEEE international symposium on biomedical imaging from nano to macro, 2007, pp 780–783

38. Zhang J, Ma K, Er M, Chong V (2004) Tumor segmentation from magnetic resonance imaging by learning via one-class support vector machine. In: International workshop on advanced image technology, pp 207–211

39. Zuiderveld K (1994) Contrast limited adaptive histogram equalization. In: Graphics gems IV, Academic Press, Waltham, pp 474–485