HHS Public Access

Author manuscript

IEEE Trans Med Imaging. Author manuscript; available in PMC 2015 March 24.

Published in final edited form as:

IEEE Trans Med Imaging. 2012 October ; 31(10): 1941–1954. doi:10.1109/TMI.2012.2210558.

GLISTR: Glioma Image Segmentation and Registration

Ali Gooya [Member, IEEE],
Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Iran

Kilian M. Pohl [Member, IEEE],
Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104 USA

Michel Bilello,
Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104 USA

Luigi Cirillo,
Bellaria Hospital, University of Bologna, 40139 Bologna, Italy

George Biros [Member, IEEE],
Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, TX 78712 USA

Elias R. Melhem, and
Department of Radiology, Hospital University of Pennsylvania, Philadelphia, PA 19104 USA

Christos Davatzikos [Senior Member, IEEE]
Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104 USA

Ali Gooya: a.gooya@modares.ac.ir

Abstract

We present a generative approach for simultaneously registering a probabilistic atlas of a healthy population to brain magnetic resonance (MR) scans showing glioma and segmenting the scans into tumor as well as healthy tissue labels. The proposed method is based on the expectation maximization (EM) algorithm that incorporates a glioma growth model for atlas seeding, a process which modifies the original atlas into one with tumor and edema adapted to best match a given set of patient's images. The modified atlas is registered into the patient space and utilized for estimating the posterior probabilities of various tissue labels. EM iteratively refines the estimates of the posterior probabilities of tissue labels, the deformation field and the tumor growth model parameters. Hence, in addition to segmentation, the proposed method results in atlas registration and a low-dimensional description of the patient scans through estimation of tumor model parameters. We validate the method by automatically segmenting 10 MR scans and comparing the results to those produced by clinical experts and two state-of-the-art methods. The resulting

Correspondence to: Ali Gooya, a.gooya@modares.ac.ir.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

segmentations of tumor and edema outperform the results of the reference methods, and achieve a similar accuracy from a second human rater. We additionally apply the method to 122 patients scans and report the estimated tumor model parameters and their relations with segmentation and registration results. Based on the results from this patient population, we construct a statistical atlas of the glioma by inverting the estimated deformation fields to warp the tumor segmentations of patients scans into a common space.

Index Terms

Diffusion-reaction model; expectation maximization (EM) algorithm; glioma atlas; joint segmentation-registration

I. Introduction

Glioma is a class of brain tumors, primarily originating from glial cells in the brain. The most common malignant form of glioma is Gliobalstoma (GBM) that exceeds in its mortality rate beyond other types of brain tumors [1]. Despite of significant advances in imaging, introduction of novel chemotherapeutic agents, and technological development, the median life expectancy of patients with GBM is less than 12 months, with only 5% of these patients surviving five years after diagnosis [2]. Quantifying the volume and other morphological characteristics of a tumor is an important indicator of the disease progression in retrospective studies such as [3], as well as the efficacy of the therapy [4]. Conventionally, such assessments are often done using so-called McDonald criteria [5], a 2-D evaluation of the largest tumor diameters. This technique, however, falls short in accurate quantification of the tumor volume due to its 2-D description of the tumor structure. The central concept for this quantification is segmenting the whole pathology into various underlying components from magnetic resonance images (MRI). Usually automatic segmentation is more reproducible and therefore preferable over manual delineations. However, the automatic segmentation of GBMs is considered very challenging as the pathology is highly heterogeneous and may consist of several components including necrosis (dead central part), enhancing tumor ring and edema (swelling). Moreover gliomas are highly variable in their size as well as shape and might cause deformation of surrounding tissues (mass-effect). Part of the presented work is a new automatic segmentation approach for this task.

There have been a variety of publications discussing automatic segmentation of brain tumors. Early works including the fuzzy clustering method of [6] revealed the intensity overlap of tumor with other tissues. Clark *et al.* [7] applied knowledge-based techniques to design more sophisticated features using multi-modality MR images for fuzzy clustering. In a similar framework, [8] extracted the 3-D connected components of the segmented tumor voxels to eliminate the false positives. The limited success of these methods on the larger number of data sets proved the need for more sophisticated approaches, which were based on the level sets and active contours [9]–[11]. In these methods, the spatial smoothness of segmented tumor is based on the curvature smoothing term that reduces the risk of curve leaking to the background. However, an excessively large curvature weight could also

prevent the full segmentation of tumor. To facilitate specifying this free parameter, Lefhon *et al.* [9] used GPU programming which enabled the user to chose it interactively, a process which otherwise can be a tedious task. Furthermore, curvature smoothing has a global effect and can not be locally adaptive.

Another series of tumor segmentation methods are based on the modern pattern classification techniques. The majority of these methods can be broadly categorized into discriminative [12]–[15], and generative approaches such as [16]–[24]. Generative approaches explicitly define a model for the joint probability distribution of voxel labels (target variables) and intensities (observed variables). Often they compute the product of the class conditional (likelihood) and prior probability functions of the labels, specifying a full probabilistic model which can be used to simulate (*i.e., generate*) all variables. On the other side, discriminative methods *only* define a model to infer the voxel labels conditioned on their intensities, without assuming a meaningful prior probability.

Support vector machine (SVM) classifiers [13]–[15] comprise a subgroup of discriminative methods which their label inferring models are SVM scoring functions. These functions are trained by maximizing the margin in the training and minimizing expected error in the testing data. The underlying assumption in this training is that the all voxels are independent and identically distributed (i.i.d.) random variables. However, in an actual image due to the obvious spatial dependencies of the neighboring voxels, the i.i.d assumption does not hold true. Hence, the quality of the final segmentation is significantly degraded. To eliminate these errors Gaussian smoothing and inclusion of neighboring intensities in the feature vectors are often considered. Similarly, Lee *et al.* [12] proposed pseudo-conditional random fields where the spatial interaction of the neighboring voxels is directly modeled on the posterior probabilities, through a potential function which depends on both voxel labels and intensities. Yet, his work is another form of discriminative methods because it does not specify an explicit model for prior probabilities of the labels. A more systematic approach to model the spatial smoothness of the labels can be achieved through generative models discussed below.

Generative methods can be further subcategorized according to their specific ways to determine the likelihood and the prior probability functions. For instance, Prastawa *et al.* [17] first affinely registered a spatially smooth probabilistic atlas to the patient scans to determine the prior probabilities of healthy tissues. Next, they refined a training set of healthy tissue voxels and used a Parzen kernel approach [25] to nonparametrically infer the likelihood probabilities from training samples. Similarly, [18] used the probabilistic boosting trees [26] to infer the likelihood probabilities from a training set, and a Markov random field (MRF) [25] as the prior probabilities of the labels. Corso *et al.* [23] computed the likelihoods from a Gaussian mixture models [25] and used a graph structure [27], [28] to implicitly model the prior probabilities of the labels in the from of longer dependencies induced by the graph.

One advantage of generative methods over discriminative ones, lies in the fact that they provide a systematic framework to explicitly model the spatial smoothness through prior probabilities of the voxel labels. To achieve this, using the Bayes rule the posterior

probability of the labels in the whole image is computed and subsequently maximized in a process which is called *maximum a posterior* (MAP) estimation. Since the posterior probability is proportional to both likelihood and prior probabilities, the MAP framework results in a segmentation which is in general constrained by both observed intensities and spatial smoothness implied by the priors. The maximization itself can be achieved by the graph cut methods [18]–[21], [23], expectation-maximization (EM) algorithm [22], [24] and level set methods [17].

In spite of relative success of generative methods, the applied prior probabilities for the spatial integrity of tumor labels has been limited to various forms of random fields such as Markov random fields (MRFs) [18], [24], discriminative random fields [19] or graph hierarchical dependencies [23]. These priors in general might lead to spatially smooth segmentation, however, the effect of smoothness prior is globally determined by a temperature parameter, a scalar which is involved in definition of the clique potentials [29]. We are ideally interested in a set of probability functions, that are spatially variable and adaptive to the local structure of the brain. Such locally variable prior probabilities for *healthy tissues* in normal looking brain segmentation can be achieved through registration to probabilistic atlases of a healthy population [30]–[32]. These atlases are defined as a set of spatially variable and smooth probability maps for white matter, gray matter and cerebrospinal fluid [see Fig. 1(a)–(c)]. However, inferring such a probability map for tumor is far more difficult because the tumor distribution in the brain is specific to the patient.

To infer such a subject specific tumor probability map, in a multiple MR modality study, Prastawa *et al.* [22] modified a probabilistic atlases of a healthy population into one with tumor. They modeled the difference of T1 and T1-CE (contrast enhanced) modalities as a mixture of gamma (tumor) and Gaussians (noise) distributions. Next, they compute the posterior probability of the gamma distribution and *affinely* register it into the probabilistic atlas of healthy population. The affinly registered posterior is used as the tumor prior probability map to mask out the original probability maps in the atlas. One major issue is that the method relies on the contrast enhancement signal hence necrotic cores are not easily detected. Moreover, since only an affine registration is used, the mapping between the probabilistic atlas and patient scans is structurally inconsistent, an issue which leads to misclassification of the tissue labels.

Our method of inferring such a tumor probability map in the probabilistic atlas is based on the tumor growth and infiltration model by [33]. The method has several advantageous over previous work in [22]. First, the model mimics the underlying biophysical process of a tumor growth in the brain hence the generated tumor map provides more realistic information about the tumor distribution in the brain than [22]. Second, our approach does not rely on contrast enhancement signal alone and handles necrotic cores. Finally, the tumor model reduces the structural inconsistency between the atlas and patient's images by adapting a mass-effect on the atlas. The remaining inconsistency with patient scans is captured by a deformable registration model which in turn improves result of the segmentation.

In this paper, we present a generative method for glioma segmentation, called Glioma image segmentation and registration (GLISTR), based on the EM algorithm which jointly segments and registers the patient scans to a probabilistic atlas of healthy individuals. The main contribution of the paper is the incorporation of the diffusion-reaction tumor growth model [33] to infer a patient specific atlas and subsequently use it as a set of probability maps to compute the posterior probabilities of tissue labels. In other words, we rely on the tumor growth model to deform and mask out the original atlas with a synthetically generated tumor probability map, a process which we call it atlas seeding. The seeded atlas is registered to patient scans and guides the segmentation process. EM iteratively refines the estimates of the posterior probabilities of tissue labels, the deformation field and the tumor growth model parameters. Therefore in addition to segmentation, the proposed method results in both atlas normalization and a low dimensional description of the patient tumor by estimation of the growth model parameters. These quantities are relevant for the prognosis of treatment or estimation of the patient survival rates [34]. These additional parameters, however, greatly increase the complexity of the optimization problem compared to [30]–[32] which are targeted to healthy brain segmentation.

We previously focused just on registering brain scans of a healthy individual to a tumor patient in [35]; an issue that was also investigated in similar registration works such as [36], [37]. Unlike the approach of this paper, the method proposed in [35] required the segmentations of both brain images to initialize the registration process. Our proposed work, however, is an extended version of [38]. We additionally provide a detailed derivation and validate the method by comparing the segmentation accuracies to the discriminative [14] and generative [22] methods. These comparisons are based on the reference volumes from two clinical experts in our institute. Moreover, we apply the proposed method (GLISTR) to 122 data sets, construct a statistical glioma atlas based on those results, and show the distribution of the extracted average tumor probability and the estimated tumor seeds in the brain.

The remainder of the paper is organized as follows. In Section II we review the diffusion-reaction model and describe the construction of the atlas. The atlas is then used in Section III to guide our joint registration and segmentation framework. In Section IV we present our quantitative evaluation and results of the application to sample patients and conclude the paper in Section V.

II. Atlas Generation

In this section, we first briefly review the tumor model and describe the procedure we take to the convert the original atlas of the healthy population into the one with tumor and edema. We define the atlas as a set of probability maps that specify the spatial distribution of brain tissues [see Fig. 1(a)–(c)]. The modified atlas then guides the EM algorithm in registering and segmenting of subjects with glioma.

Our tumor growth model [33] is a framework for modeling gliomas growth and their mechanical impact on the surrounding brain tissue (the so-called, mass-effect). It is an Eulerian continuum approach that results in a strongly coupled system of nonlinear partial

differential equations (PDEs): a reaction-diffusion model for the tumor growth [see (1)] and a piecewise linearly elastic material for the brain tissue [see (2)]. These equations are solved using fictitious domain method [39], where the target brain domain is embedded in a larger rectangular space which encompasses the brain background. The reaction-diffusion equation determines the rate of change in the tumor probability, as the tumor cells diffuse into brain and cause displacement (advection) and further proliferation (reaction).Quite comprehensive tumor models based on anatomical MRI and diffusion tensor images (DTI) have recently been proposed by [40]–[42]. In this paper to circumvent the need for DTI, we choose the model by [33], which only requires anatomical MRI. The good results of Section IV seem to indicate that this was a reasonable choice as our approach reliably registers and segments the scans even without the additional information provided by DTI. Based on this model, we artificially seed a tumor in the healthy atlas and grow it.

To define this model formally, let Ω_A denote the space of atlas and $[0, T]$ a time interval for growing the tumor. The evolution of tumor probability $\pi_{TU}(\mathbf{x}, t)$ in $\Omega_A \times [0, T]$ is determined by the following diffusion-reaction model [(1)] and the mass effect model [(2)]:

$$\underbrace{\frac{\partial \pi_{TU}}{\partial t}}_{\text{Rate of change}} - \underbrace{\nabla(D\nabla\pi_{TU})}_{\text{Diffusion}} + \underbrace{\nabla(\pi_{TU}\mathbf{v})}_{\text{Advection}} = \underbrace{\rho\pi_{TU}(1-\pi_{TU})}_{\text{Reaction}} \quad (1)$$

$$\underbrace{\nabla\cdot[\lambda\nabla\cdot\mathbf{u}\mathbf{I}+\mu(\nabla\mathbf{u}+\nabla\mathbf{u}^T)]}_{\text{Linear elastic differential form for displacement}} = \underbrace{p\nabla\pi_{TU}}_{\text{Displacement force}} \quad (2)$$

where ∇ and $\nabla\cdot$ are gradient and divergence operators respectively, $(\cdot)^T$ denotes the transpose operation, $\mathbf{u} : \Omega_A \rightarrow \Omega_A$ (i.e., a displacement in the atlas space) is the mass effect caused by the presence of the tumor, $\mathbf{v} = \partial\mathbf{u}/\partial t$ is the relevant velocity field, p is a scalar which determines the strength of the tumor mass effect, D is a spatially variable function capturing diffusion coefficient within white (D_{WM}) and gray matters (D_{GM}), and ρ is proliferation coefficient. We fix $\rho = 0.025$ (the default values of [33]) and optimize the model with respect to D because the profile of π_{TU} is effectively determined by the ratio of D/ρ. We also fix $D_{GM} = 1e^{-10}$ as the method is relatively insensitive to that parameter. λ and μ are Lamé coefficients that relate to stiffness (Young's module) and compressibility (Poisson ratio) [43] properties of the brain tissues. In order to minimize the computational cost of the estimation problem, we fix those to the preset values utilized in [33] and [44], [45], i.e., $\lambda = 6500$, $\mu = 725.0$ for brain parenchyma, and $\lambda = 57.0$, $\mu = 227.0$ for ventricles.

Now, if we denote with $\mathbf{x}_0$ the initial seed location of the tumor and with d the voxel size of the Ω_A, then our tumor growth model is completely defined by the parameters $\mathbf{q} \equiv \{\mathbf{x}_0, p, D_{WM}, T\}$ given the initial conditions for the $\mathbf{u}(\mathbf{x}|\mathbf{q}, t = 0) = 0$ and $\pi_{TU}(\mathbf{x}|\mathbf{q}, t = 0) \simeq exp(-(\mathbf{x} - \mathbf{x_0})^2/d^2)$. This approximation is because the Gaussian form is only computed for the seed and its neighboring voxels. We assign $\pi_{TU}(\mathbf{x}|\mathbf{q}, t = 0)$ to zero for other voxels to constraint the initial tumor probability into the local proximity of the seed.

Once we solve the above equation for **u** and π_{TU}, we combine those results at $t = T$ with the original atlas of healthy brains $\pi_X^o(\cdot)$ to infer tissue probability maps $\pi_X(\cdot|\mathbf{q})$ for tissue X. To simplify notation, we omit $t = T$ from **u** and π_{TU} and simply denote with $\pi_{TU}(\mathbf{x}|\mathbf{q})$ the spatial probability map of glioma being present at location $x \in \Omega_A$ at time T and $\mathbf{u}(x|\mathbf{q})$ the corresponding mass effect at that location and time.

We initially note that a typical tumor may consist of both enhancing (EN) and necrotic (NE) regions which have different imaging specifications. We address this issue in our framework by characterizing them with distinct intensity distributions. However, our biophysical model views them as one class generating the spatial probability map π_{TU} for both pathologies. To resolve this issue we further assume that both NE and EN labels are equally probable. Thus, to infer their corresponding probability maps, we simply divide π_{TU} by two

$$\pi_{EN}(\mathbf{x}|\mathbf{q}) = \pi_{NE}(\mathbf{x}|\mathbf{q}) \equiv 0.5\pi_{TU}(\mathbf{x}|\mathbf{q}). \quad (3)$$

We then construct $\pi_X(\cdot|\mathbf{q})$ for GM and CSF by deforming the corresponding spatial probabilities $\pi_X^o(\cdot)$ of the healthy population via the mass-effect **u** and weighing them with $(1 - \pi_{TU})$

$$\pi_{GM}(\mathbf{x}|\mathbf{q}) \equiv \pi_{GM}^o(\mathbf{u}(\mathbf{x})) \cdot (1 - \pi_{TU}(\mathbf{x}))$$

$$\pi_{CSF}(\mathbf{x}|\mathbf{q}) \equiv \pi_{CSF}^o(\mathbf{u}(\mathbf{x})) \cdot (1 - \pi_{TU}(\mathbf{x})). \quad (4)$$

We construct the spatial probability map of edema π_{ED} based on the assumption that edema is in close proximity of the tumor, which we model via the Heaviside function defined as

$$H(a) = \begin{cases} 0, & \text{for} a \leq 0 \\ 1, & \text{otherwise} \end{cases} \quad (5)$$

and it is confined to the mass deformed white matter of the healthy brain, which we model with $\pi_{WM}^o(\mathbf{u}(\mathbf{x}))$. Thus

$$\pi_{ED}(\mathbf{x}|\mathbf{q}) \equiv 0.5\pi_{WM}^o(\mathbf{u}(\mathbf{x})) \cdot (1 - \pi_{TU}(\mathbf{x})) \cdot H(\pi_{TU}(\mathbf{x})) \quad (6)$$

where the multiplication by 0.5 essentially means that no particular preference is given to either of edema or white matter labels; the tissue type is to be determined by the image intensities. We finally define the subject specific spatial probability map of the white matter as the remaining probability given the probabilities of the other labels

$$\pi_{WM}(\mathbf{x}|\mathbf{q}) \equiv 1 - [\pi_{TU}(\mathbf{x}|\mathbf{q}) + \pi_{ED}(\mathbf{x}|\mathbf{q}) + \pi_{CSF}(\mathbf{x}|\mathbf{q}) + \pi_{GM}(\mathbf{x}|\mathbf{q})]. \quad (7)$$

A sample set of the generated probability maps $\pi_{TU}(\cdot|\mathbf{q})$, $\pi_{ED}(\cdot|\mathbf{q})$, $\pi_{CSF}(\cdot|\mathbf{q})$, $\pi_{GM}(\cdot|\mathbf{q})$ and $\pi_{WM}(\cdot|\mathbf{q})$ is shown in Fig. 1(e)–(i), illustrating the impact of the mass effect and tumor invasion in originally healthy atlas. In the rest of paper, we simplify our notation by

denoting the probability maps generated in this section with $\pi_k(\cdot|\mathbf{q})$, $1 \leq k \leq K = 6$, with k enumerating the set of our segmented labels: {*NE, EN, GM, CSF, ED, WM*}.

III. Joint Segmentation-Registration

We now describe the framework for joint segmentation-registration which is guided by the atlas defined in the previous section. We assume that a set of J co-registered, inhomogeneity-corrected, and skull stripped MR images is given in the reference (fixed) domain Ω_F so that for any sample voxel $\mathbf{x} \in \Omega_F$, $\mathbf{y}(\mathbf{x}) \equiv [y_1(\mathbf{x}), \ldots, y_J(\mathbf{x})]^T$ is an independent observation vector that corresponds to the J image modalities. We then define observation set as: $\mathbf{Y} \equiv \{\mathbf{y}(\mathbf{x})|\mathbf{x} \in \Omega_F\}$. The goal of this section is to derive an algorithm for estimating the intensity distributions of each structure $\mathbf{\Phi}$, the atlas coefficient $\mathbf{q}$, and the deformation between the atlas and the reference domain $\mathbf{h}$.

We further specify $\mathbf{\Phi}$ by assuming the conditional probability distribution function (pdf) of each $\mathbf{y}(\mathbf{x})$ is a weighted mixture of K Gaussians

$$f(\boldsymbol{Y}|\mathbf{\Phi}, \mathbf{h}, \mathbf{q}, \mathbf{x}) \equiv \sum_{k=1}^{K} \pi_k(\mathbf{h}(\mathbf{x})|\mathbf{q}) f_k(\mathbf{y}(\mathbf{x})|\mathbf{\Phi}) \quad (8)$$

where $f_k \sim N(\mathbf{m}_k, \mathbf{\Sigma}_k)$ is a multivariate Gaussian distribution with mean $\mathbf{m}_k$ and the covariance matrix $\mathbf{\Sigma}_k$, and $\mathbf{\Phi} \equiv \{\mathbf{m}_1, \ldots, \mathbf{m}_K, \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_K\}$. The mixture weights are determined by $\pi_k(\mathbf{h}(\mathbf{x})|\mathbf{q})$ which are originally defined in the atlas space Ω_A (see Section II) and registered to the patient space through $\mathbf{h} : \Omega_F \rightarrow \Omega_A$, a vector field mapping the patient space into the atlas. Based upon these assumptions, we write the likelihood of $\mathbf{Y}$ as

$$f(\boldsymbol{Y}|\mathbf{\Phi}, \mathbf{h}, \mathbf{q}) = \prod_{\mathbf{x} \in \Omega_F} f(\boldsymbol{Y}|\mathbf{\Phi}, \mathbf{h}, \mathbf{q}, \mathbf{x}). \quad (9)$$

Our problem of joint segmentation, registration and atlas parameter estimation can be defined as the solution of the following:

$$(\mathbf{\Phi}_o, \mathbf{h}_o, \mathbf{q}_o) \equiv \underset{\mathbf{\Phi}, \mathbf{q}, \mathbf{h}}{\operatorname{argmax}} \log f(\boldsymbol{Y}|\mathbf{\Phi}, \mathbf{h}, \mathbf{q}). \quad (10)$$

One way of computing the solution to this problem is via an implementation of the EM algorithm [46]. EM is an iterative algorithm which instead of solving (10), it maximizes the expectation of the log-likelihood of the complete observation under the current estimates of posteriors [47]. In other words, at every iteration given the last estimate of the unknown parameters $\mathbf{\Phi}'$, $\mathbf{h}'$ and $\mathbf{q}'$, it sequentially maximizes the following statement:

$$Q(\mathbf{\Phi}, \mathbf{h}, \mathbf{q}|\mathbf{\Phi}', \mathbf{h}', \mathbf{q}') \equiv \sum_{\mathbf{x} \in \Omega_F} \sum_{k=1}^{K} p_k(\mathbf{x}) \log(\pi_k(\mathbf{h}(\mathbf{x})|\mathbf{q}) f_k(\mathbf{y}(\mathbf{x})|\mathbf{\Phi})). \quad (11)$$

where $p_k(\mathbf{x})$ stands for the posterior probability of class k at voxel $\mathbf{x}$. The structure of the proposed EM algorithm consists of iterations between the E-Step and M-Step, during which

the p_k (**x**) and parameters {**Φ**, **h**, **q**} are respectively updated. Further details are explained in the rest of this section.

A. E-Step

We compute the posterior probabilities given the current estimate of the parameter

$$p_k(\mathbf{x})=\frac{f_k(\mathbf{y}(\mathbf{x})|\mathbf{\Phi}')\pi_k(\mathbf{h}'(\mathbf{x})|\mathbf{q}')}{\sum_{l=1}^{K} f_l(\mathbf{y}(\mathbf{x})|\mathbf{\Phi}')\pi_l(\mathbf{h}'(\mathbf{x})|\mathbf{q}')}. \quad (12)$$

B. M-Step

We update the parameters **Φ**, **h** and **q**. The mean and covariances of the Gaussian distributions in **Φ** are updated using the following equations such as in [32]

$$\mathbf{m}_k' \leftarrow \frac{\sum_{\mathbf{x}\in\mathbf{\Omega}_F} p_k(\mathbf{x})\mathbf{y}(\mathbf{x})}{\sum_{\mathbf{x}\in\mathbf{\Omega}_F} p_k(\mathbf{x})} \quad (13)$$

$$\mathbf{\Sigma}_k' \leftarrow \frac{\sum_{\mathbf{x}\in\mathbf{\Omega}_F} p_k(\mathbf{x})[\mathbf{y}(\mathbf{x})-\mathbf{m}_k'][\mathbf{y}(\mathbf{x})-\mathbf{m}_k']^t}{\sum_{\mathbf{x}\in\mathbf{\Omega}_F} p_k(\mathbf{x})}. \quad (14)$$

To update the deformation field **h** we follow a variational framework which computes the differential of (11) with respect to an infinitely small test function **v**. The maximizing direction **v** is specified by equating the differential to zero

$$\lim_{\|\mathbf{v}\|\to 0}\{Q(\mathbf{\Phi}',\mathbf{h}'+\mathbf{v},\mathbf{q}'|\mathbf{\Phi}',\mathbf{h}',\mathbf{q}')-Q(\mathbf{\Phi}',\mathbf{h}',\mathbf{q}'|\mathbf{\Phi}',\mathbf{h}',\mathbf{q}')\}=\sum_{\mathbf{x}\in\Omega_F}\mathbf{v}^t(\mathbf{x}).\{\mathbf{r}(\mathbf{x})+\mathbf{W}(\mathbf{x})\cdot\mathbf{v}(\mathbf{x})\}=0. \quad (15)$$

In this equation, the gradient vector **r**(**x**) and the matrix **W**(**x**) are defined as

$$\mathbf{r}(\mathbf{x})=2\sum_{k=1}^{K} p_k(\mathbf{x})\frac{\nabla\pi_k(\mathbf{h}'(\mathbf{x})|\mathbf{q}')}{\pi_k(\mathbf{h}'(\mathbf{x})|\mathbf{q}')} \quad (16)$$

$$\mathbf{W}(\mathbf{x})=\sum_{k=1}^{K} p_k(\mathbf{x})\cdot\left[\frac{\mathbf{H}(\pi_k(\mathbf{h}'(\mathbf{x})|\mathbf{q}'))}{\pi_k(\mathbf{h}'(\mathbf{x})|\mathbf{q}')}-\frac{\nabla\pi_k(\mathbf{h}'(\mathbf{x})|\mathbf{q}')(\nabla\pi_k(\mathbf{h}'(\mathbf{x})|\mathbf{q}'))^t}{(\pi_k(\mathbf{h}'(\mathbf{x})|\mathbf{q}'))^2}\right] \quad (17)$$

where **H** is the Hessian matrix (The detailed derivations can be found in the Appendix). Equation (15) leads to **r**(**x**) + **W**(**x**).**v**(**x**) = 0, hence **v** can be obtained as

$$\mathbf{v}(\mathbf{x})=\mathbf{W}^{-1}(\mathbf{x})\cdot\mathbf{r}(\mathbf{x}). \quad (18)$$

Because the matrix **W** can be singular, we add an identity matrix component to **W** for damping, therefore the update of **h**(**x**) can be written as

$$\mathbf{h}'(\mathbf{x}) \leftarrow \mathbf{h}'(\mathbf{x}) - [\mathbf{W}(\mathbf{x}) + c\mathbf{I}]^{-1}\mathbf{r}(\mathbf{x}) \quad (19)$$

where **I** is the identity matrix and c is the damping parameter. In this paper, we found $c = 0.1$ to produce a robust and reasonable deformation field. Notice, the update equation is computed independently at every voxel, which in general results in a nonsmooth deformation field. In order to apply the smoothness constraint, similar to Thirions' demons framework [48], we diffuse the estimated deformation vectors by a Gaussian convolution filter. We heuristically set the standard deviation to 2.

The other set of parameters updated in the M-Step are the tumor growth parameters **q**. Since no analytical expression for the derivatives of $Q(\mathbf{\Phi}', \mathbf{h}', \mathbf{q}|\mathbf{\Phi}', \mathbf{h}', \mathbf{q}')$ with respect to **q** exists, we maximize (11) using a derivative free pattern search library [49]. Since this operation is computationally expensive it is performed only after having an adequate convergence on **h** and **Φ**, otherwise we keep it fixed.

Hence, our EM algorithm is an iterative procedure; having the current estimates $\mathbf{\Phi}'$ and $\mathbf{h}'$, we proceed by optimizing **q** using the numerical scheme in [49]. Next, with $\mathbf{q}'$ fixed in the optimal point, we iteratively update the parameters Φ and **h**. This requires an iterative sweeping through equations in (12), (13), (14), and (19) until a convergence is achieved. Then, the whole procedure starts over with newly updated parameters. Typical examples, each with three iterations of this procedure, are shown in Fig. 2. This completes our derivation of the EM implementation for joint segmentation and registration.

IV. Experiments

In this section, we first describe our data preprocessing pipeline, and introduce organization of our experiments which mainly cover our validation framework and application of the method to a large data set of glioma patients. Finally, we present the results which include quantitative and qualitative evaluations of the segmentations and registrations, estimated tumor model parameter values and a statistical glioma atlas obtained using the proposed method.

A. Data Specification and Preprocessing

Four MR modalities (T1/T2/FLAIR/T1-CE) were acquired for every patient in our data sets. Following a clinical protocol in University of Pennsylvania, all patients were imaged on a 3.0T MRI scanner system (MAGNETOM Trio Timstem, Siemens Medical Systems, Erlangen, Germany) with the same protocol: magnetization prepared rapid acquisition gradient echo 760/3.1/950, matrix size (192 × 256), pixel spacing (0.9766 × 0.9766 mm. A slice thickness of 3 mm was used to acquire FLAIR and T1-CE modalities. Also, T1 and T2 modalities were acquired with slice thickness of 0.9 and 1 mm, respectively.

The preprocessing starts with image inhomogeneity correction [50] and proceeds with skull-striping and cerebellum removal of all modalities [51]. Next, these images are co-registered

using an affine registration algorithm [52]. Finally, the co-registered images are subsequently affine registered to the atlas space using the same algorithm in the previous step. This results in a set of scans with dimensions of (256 × 256 × 128) and the voxel size of (0.9375 × 0.9375 × 1.5 mm^3).

To speed up the simulations of the tumor growth, we solve the partial differential equations in (1) on a lattice of (65 × 65 × 65) nodes that is down sampled from the original atlas. The computed tumor probability map is then up sampled to the original resolution and used to mask out the healthy atlas using equations in (4), (6), and (7).

B. Organization of the Experiments

We first study the sensitivity of the method with regard to initial seed location and demonstrate its robustness to changes in this parameter. We present these results in terms of the quality of acquired segmentations versus the converged cost values and show that a higher cost value produces lowered quality results.

Next we compare our segmentation results with two state-of-the-art methods and a human expert rater. The first method [22] is a generative method based on EM algorithm which has two major distinctions with GLISTR: it uses an affine model for registration, and it does not rely on a tumor model to alter the atlas but constructs a probability model of the tumor purely based on intensity analysis. To differentiate GLISTR from discriminative segmentation methods, we also compare the results of GLISTR to a second method denoted as SVM [14]. In our implementation, the intensities of co-registered FLAIR, T2, T1 and T1CE images specify the feature vector for every voxel. The target tissue labels of the SVM segmentor are the same labels that are segmented by GLISTR, i.e., necrosis (NE) and enhancing tumor (NE), edema (ED), cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM). For each patient, an expert choses small sample ROIs inside the target tissue regions, that we use to train a multi-class SVM model [53] and segment the whole brain.

We measure the overlap of the reference (R) and the segmented (S) volumes using [54]

$$\text{Dice Score} = \frac{2|R \cap S|}{|R| + |S|} \quad (20)$$

where $|x|$ denotes the cardinality of the set x. We use manual segmentations as for reference volumes in our measurements. Furthermore, because we did not have separate reference segmentations for necrosis and enhancing labels, we unify the NE and EN labels to make a single tumor label (TU)

$$TU = NE \cup EN. \quad (21)$$

We then evaluate the Dice ratios for this combined label for which we have the reference maps from our clinicians.

The manual segmentations for our experiments were made by two specialists who independently segmented 10 patients scans into the set of target tissue types. To reduce the amount of required labor, only five cases were fully segmented into edema and tumor labels.

For the other cases, every third slice of the pathology was considered for tumor and edema delineations. To ensure that there is no significant difference across the Dice scores measured using these selected slices and those measured based on entire volumes, we additionally computed the Dice scores with respect to every third slice in the fully segmented five cases. Then we computed the difference between these scores and those obtained based on entire volumes for all cases. The average of these five differences was less than 0.75%.

For tumor and edema labels, we also evaluate the segmentation accuracies in terms of the surface distances between the reference and the acquired segmented volumes with various methods.

Moreover, for CSF, gray and white matter tissues, similar to the validation method taken by [55], these manual segmentations were limited into small rectangular ROIs with the size of (80×80) pixels located in the healthy part of brain. The availability of two expert segmentations allowed us to chose one set as the reference and compare the accuracy of GLISTR versus the second expert.

In the second set of experiments, we estimate the tumor model parameters and present the relations between these parameters and results of segmentation-registration for 122 tumor patients. These data sets consist of of various brain neoplasms including: 74 cases of Glioblastoma World Health Organization (WHO) grades of III/IV, 11 cases of Metastasis, 5 Oligodendrogliomas, 4 Anaplastic astrocytomas WHO grades of II/IV, and 32 cases of other low grade gliomas. The age of these patients ranged from 22 to 87 with the average 56.72 and the standard deviation of 14.67 years.

We should point out that the metastatic brain tumors are not glial based as they originate from a primary cancer in other parts of the body. These tumors are usually noninfiltrative [56], hence the diffusion-reaction model which explicitly formulates infiltration, may not be an appropriate model for their growth. Nevertheless, since the manual segmentation of these lesions is difficult (mostly due to tissue heterogeneity) we have included them in our studies to show the versatility of the proposed method.

C. Sensitivity with Respect to Initial Seed Location

In this section we demonstrate the sensitivity of the proposed method with respect to initial tumor seed location. In principle, this location is initially determined by the user clicking an approximate center of tumor. The seed location is then refined in the M-step as a part of tumor parameters ($\mathbf{q}$). We show the robustness of GLISTR in producing satisfying results when this initial seed location is displaced.

To achieve this objective we chose a patient with an approximate tumor size of 5 cm^3 [Fig. 3(a) and (b)]. Then around the tumor center, we pick up eight different initial seed locations placed 2 ~ 3 cm apart from each other. Given the fact the size of the search box for the tumor seed is 3 cm^3, the entire tumor region is examined. We then run GLISTR for each of these selected seed points and compare the results.

Fig. 2(a) shows the optimization process and the converged cost value for each seed. Note that the largest drop of the cost corresponds to the first round of convergence in **h**. This brings the probabilistic atlas into a better correspondence with patient anatomy compared to the starting point of optimization (with affine registered atlas). Hence, the subsequent optimization of **q** generally have much lower cost values. As seen in Fig. 2(b), converged local minimums are quite similar, with the lowest (best) and highest minimums found in S.7 and S.4, respectively. We quantitatively compare and visually inspect the various segmentations qualities obtained from these experiments.

For this purpose we first use the Dice ratio introduced in (20) and measure the overlaps of the segmented ED and TU labels with expert provided reference volumes. Table I summarizes these measurements. We note that except for the case S.4, which produces significantly smaller Dice ratios, all other overlaps vary minimally across different seeds. Moreover, we also note that the particular case of S.4. the corresponding cost value is significantly larger, showing that the quality of results is sensitive to the converged cost value and degrades with increase in the latter.

Fig. 3(d)–(k) show the corresponding segmentations within the slice through the largest section of the tumor. With an exception of Fig. 3(g), we observe that these segmentations visually resemble the expert provided reference segmentations of ED and TU labels in Fig. 3(c) with a high degree. The particular case shown in (g) corresponds to case S.4, for which we measured the largest converged cost as well as the worst Dice ratios.

In summary, these observations indicate that the cost function has multiple local minimums which are associated with the converged values in Table I. For 7 out of 8 initial seed locations, the converged local minimums produced both visually and quantitatively satisfying results (except for S.4). This shows that the GLISTR is reasonably robust to changes in the initial seed location.

D. Segmentation and Registration Results

In this section, we first visually compare the segmentation results of GLISTR with those obtained from [22], SVM and a second clinical expert in Fig. 4. This is followed by our quantitative evaluations of the segmentation accuracies in Fig. 5. Finally, for visual assessments of the *joint* segmentation and registration efficacy of GLISTR, we present sample results for ten patients in Fig. 6.

Fig. 4 shows sample segmentation for five patients. As seen, the discrepancy in the results from two experts is noticeable (e.g., refer (a) and (b) to the forth row). This discrepancy is because of ambiguity of the intensities in the pathological region. Subjective interpretation of this ambiguity results in different expert segmentations. Moreover, by assuming the segmentations of the first rater as reference volumes, we observe a very good agreement between our results with the reference segmentations.

Notice that using [22] in the column (d) of Fig. 4, many necrotic parts of the tumors are miss-classified as CSF (red in color). This is because [22] subtracts T1CE and T1 modalities to form a prior map for the pathology. Since those modalities show a same range of

intensities in necrosis, their difference is very small and hard to detect. This leads to a poor estimation of the prior map for the necrosis, hence its segmentation fails. In contrary, GLISTR does not segment the necrosis independently but as a part of tumor whose probability map in the atlas is generated by a bio-mechanical model. Hence, it produces better results in all cases. On the other hand, SVM segmentations in column (e) seem to segment the necrosis and most of the tumors correctly. However, it also detects many tumor false-positives in the rest of the brain. This is because this type of segmentation is purely voxel-based and does not rely on any spatial constraint. Whereas in GLISTR, we effectively constrain tumor detection into the pathological region where the model generated tumor probability map has nonzero values. These observations are further verified by our numerical evaluations summarized in Fig. 5. As seen, results of our proposed method are comparable to the second rater and outperforms [22] and SVM on tumor and edema. We verify those findings by performing statistical hypothesis paired t-tests between our computed Dice scores and those from the second expert and the other two methods in columns (b)–(d). The results from these tests confirm that segmented tumor and edema using GLISTR, are not significantly different from the second human rater segmentations and they are significantly better than the other reference methods (p-value < 0.01). Moreover, on the healthy tissue labels, all automatic methods had a similar accuracies and the difference was only significant for the second expert.

To provide a further insight on the accuracies in terms of physical units, in Table II we present the average surface distances between the reference and acquired segmented volumes in millimeters. As seen, GLISTR's performance compares favorably to the second human rater (RATER) and stays higher than the other methods. The large surface errors for SVM and [22] are due to existence of false tumors and edema positives that are scattered within the brain (e.g., refer to fifth row in Fig. 4). These regions increase the average surface distances from the reference segmentations that are limited to the borders of the pathologies.

The joint segmentation and registration results of ten patients are given in Fig. 6. As seen, the segmentations obtained using GLISTR in column (d) show a high visual correspondence with patients anatomies. Moreover, it is interesting to observe that the registered atlas probability maps in column (e) and (f) closely match the patient segmented labels, which indicate good registration as well.

Finally we comment on the computation time required by each method. Depending on the size of the tumor and complexity of the images, GLISTR requires 3 ~ 6 h to converge on segmentation and estimate of the tumor parameters. On the other hand, the method proposed in [22] needs shorter periods of around 30 min to estimate the labels, and SVM requires 40 ~ 50 min for training and testing. However, we note that the prolonged computation time for GLISTR is due to inherent complexity that this joint framework requires for deformable registration, segmentation and estimation of the relevant tumor parameters.

E. Relations of Tumor Parameters With Segmentation-Registration Outcomes

Our objective in this section is to investigate the relations between the estimated tumor parameters $\mathbf{q}$ and the segmentation- registration outcomes defined as: volumes of segmented tumor $Vol(TU)$, edema $Vol(ED)$ and the total average expansion (mass-effect) in the tumor

region. The latter is specified by the average of determinant of the Jacobian of the total deformation field [57] in the tumor region, i.e., $|Jac(\mathbf{h} \circ \mathbf{u})|$. In Fig. 7, we plot $Vol(TU)$, $Vol(ED)$, and $|Jac(\mathbf{h} \circ \mathbf{h})|$ versus the estimates of the parameters T, D_w, and p_1, respectively, for 122 Glioma patients described in Section IV-B. In addition, we compute their corresponding correlation coefficients and significance p-values in Table III. We use *Spearman* correlation coefficient because it is robust to outliers and captures nonlinear relationships [58].

It should be emphasized that we do not estimate the tumor parameters based on longitudinal studies, where a tumor growth pattern is traced over a time span to specify the diffusion-reaction parameters [41]. Rather in this paper, we estimated them by comparing a given patient's anatomy with a healthy population (the probabilistic atlas). Therefore, we view the estimated model parameters as a feature set that determines the subject's deviation from the atlas. Hence, the rationale of this analysis is that tumor parameters, estimated for each patient individually, can potentially be of diagnostic and prognostic value, and can potentially serve as guides for the tumor treatment procedure, including targeted radiation therapy.

Fig. 7(a) shows the distribution of the tumor growth durations T versus $Vol(TU)$, indicating a strong relationship between these variables. Moreover as seen in Table III, the p-value of having a correlation between T and $Vol(TU)$ is near zero. This means the size of the tumor has a very significant correlation with the estimated tumor growth duration such that the bigger tumors are estimated with larger values T of and vice versa.

In Fig. 7(b), we show the distribution of the estimated diffusion coefficient D_w versus $Vol(Ed)$. Although the relation between these parameters is not linear, two groups are distinguishable. The first group consists of those cases whose estimates of diffusion coefficients satisfy $-8 < \log(D_w)$. The majority of these cases also have large volumes of edema ($50 \text{ ml} < Vol(Ed)$). The second group comprises of the patients with the lowest estimate of the diffusion coefficient ($\log(D_w) = -9$) whose majority of edema volumes satisfy $Vol(Ed) < 50$ ml.[1] This implies that large estimates of diffusion coefficients are in general associated with greater edema volumes in the patient scans, a finding that could potentially be important in terms of guiding radiation therapy in edematous regions. The correlation coefficient of D_w and $Vol(Ed)$ in Table III is 0.6616 and the corresponding p-value is very small which implies a very significant correlation between these parameters. Next, we investigate the relation between the model estimated mass-effect and the total mass-effect which is measured by registration to the original healthy atlas. Notice that the original atlas is warped to the patient scans through a cascade of two deformation fields. The first deformation field is the tumor model mass-effect $\mathbf{u}$ as described in Section II. The second deformation field $\mathbf{h}$ subsequently warps the modified atlas into the patient's scans. Hence the total difference between the original atlas and the patients anatomy has a composite form of $\mathbf{h} \circ \mathbf{u}$. We show that the most of total mass-effect (expansion) is captured

[1] Ideally, by decreasing the lower end of the search span for D_w we could have separated this cluster into individuals (along with the vertical axis) with different diffusion coefficients. However, since our numerical tumor model solver was conditionally unstable for such low levels of D_w, we initially chose a rather conservative search span.

by **u** rather than **h**. This ensures that the estimated model mass-effect **u** is consistent with the amount of the mass-effect in the patient scans that is measured by registration.

In this paper, as a measure of mass-effect we compute the average of the determinant of the Jacobian ($|Jac(\cdot)|$) of the deformation field [57] in the tumor region. Values close to zero indicate strong expansion whereas values of one means no expansion. As such, the total mass effect measured by the registration, and the model estimated mass-effect are denoted by $|Jac(\mathbf{h} \circ \mathbf{u})|$ and $|Jac(\mathbf{u})|$, respectively.

In Table III, we have also measured the correlation coefficients between $|Jac(\mathbf{h} \circ \mathbf{u})|$ versus $|Jac(\mathbf{u})|$ and $|Jac(\mathbf{h})|$ in the tumor region. Notice that $|Jac(\mathbf{h} \circ \mathbf{u})|$ has a smaller p-value and a slightly higher correlation with $|Jac(\mathbf{u})|$ compared to $|Jac(\mathbf{h})|$. This means that the total mass-effect is most likely captured by model created deformation field **u** rather than **h** as expected.

We have also plotted $|Jac(\mathbf{h} \circ \mathbf{u})|$ versus p_1 in Fig. 7(c). Note that the amount of total expansion is reversely proportioned with $|Jac(\mathbf{h} \circ \mathbf{u})|$, so that the values close to 1 means no change in the volume. Consequently as we could see in this figure, the larger values of p_1 are in general associated with small values of $|Jac(\mathbf{h} \circ \mathbf{u})|$ (i.e., large expansions). Hence, the reported correlation between these values in Table III is significant but negative, which implies that estimates of p_1 are consistent with the expansions captured by the method. Also notice that because the numerical scheme of our tumor model was conditionally unstable with larger p_1 values, we chose a conservative maximum value of 8 for this parameter. As a result, some patients with significant mass-effect (small values of $|Jac(\mathbf{h} \circ \mathbf{u})|$) are estimated with the largest possible value for p_1.

F. Construction of Statistical Atlas for Gliomas

A highly unexplored question is whether the spatial extent and distribution of tumors relates to their phenotypes and clinical outcome. Exploring such questions calls for generation of statistical atlases from relatively large numbers of patients. We begun to explore this direction, by forming an atlas from our patient population.

To construct the brain tumor atlas that captures the statistics of the spatial distribution of glioma in this population, first the computed tumor posterior probability of every patient is transformed to the common atlas space Ω_A. We performed this spatial normalization by inverting the total deformation field that originally warps the common atlas to the individual patients scan. Hence the resulting inverted deformation field (i.e.,$(\mathbf{h} \circ \mathbf{u})^{-1}$) warps the patient scans into the common atlas space. Next we warp the posterior probability of the tumor using the inverted deformation field and then we compute the average of all warped posteriors over the patient population. Fig. 8 shows this atlas normalization procedure using two sample patient images. For comparison the original T1CE scans are indicated in the upper row. The inverted deformation field is applied on both the tumor posterior map and the T1CE scan, then the resulting images are overlayed together in the second row. As seen the mass effect has been relaxed in the atlas normalized images hence the tumors appear to be smaller and ventricles seem to be relaxed compared to the original patients scans in the

upper row. This is because $\mathbf{h} \circ \mathbf{u}$ is an expansive field due to the mass-effect, hence its inverse shrinks the images in the tumor region.

Fig. 9 shows the resulting tumor atlas created by the procedure explained above across 122 patients. We have overlaid the computed average tumor map on the probability map of the CSF (in the grey background) to specify its relative position with respect to the original atlas. The color bar indicates that, within our patient population, the region with the highest tumor probability (7.6%) coincides with the left temporal lobe which is associated with low-level perception, comprehension, naming and verbal memory.

We also render the spatial distribution of the estimated tumor seed locations in the atlas space Fig. 10. The estimated seed positions are overlayed as small colored spheres on the outer surface and the ventricles of the atlas. In this figure, the colors do not convey particular information about types of the tumor and are only intended for easy differentiation of the positions in different views. Although it is difficult to distinguish a particular pattern, we can notice that fewer seeds are located in the frontal and occipital lobes of the brain. We anticipate that, this outcome should be taken with care because it is known that the diffusion-reaction-advection based models do not model the initial stages of the tumor growth. Hence, the actual seed points might be different from locations shown in Fig. 10. Nevertheless, we believe that the nonuniformity of the glioma incidence in these figures may be the result of either a potential fundamental aspect of glioma developmental biology, or the possibility of selection bias rather than errors introduced by the method[2]. Many of the patients were imaged after they referral to our institution for resection. Therefore, we should anticipate that the presented atlas may have such a bias for tumor locations that can be safely resected.

V. Conclusion and Discussion

In this paper, we presented a method (GLISTR) for segmentation of gliomas in multi-modal MR images by joint registering the images to a probabilistic atlas of healthy individuals. The major contribution of the paper was the incorporation of a tumor growth model to adopt the normal atlas into the anatomy of the patient brain. This adoption required us to estimate a set of optimal parameters for the tumor model. Our motivation for this incorporation was to generate more realistic prior probabilities for the tumor and other labels. We used the EM algorithm to iteratively refine the estimates of tissue label posteriors, the registration parameters and the tumor growth model parameters. We evaluated the accuracy of the segmentations by computing Dice overlap scores with reference segmentations made by a clinical expert. We compared the segmentation accuracy of the GLISTR to a second human rater and two state-of-the-art methods using 10 patients scans. The results show that GLISTR outperforms the reference methods and achieves a similar accuracy of the second human rater in segmenting tumor and edema. However, the accuracy of the healthy tissue segmentations was not significantly different across the proposed and reference methods.

[2]We could confirm this because the segmentation results of all patients were visually inspected by our clinical expert and only a few cases were found as nonsatisfactory.

An important contribution of this work is that, in the process of estimating segmentation-registration, it also estimates the parameter set **q** of the reaction-diffusion tumor model. These parameters can potentially offer additional diagnostic and prognostic values clinically, as they might relate to the tumor's aggressiveness and extent of infiltration beyond imaging-visible tumor boundaries. Particularly interesting is the relationship between the diffusion coefficient D_w and the volume of edema, since it is known that aggressive gliomas are very likely to have infiltration into the region defined as edema via imaging measures, therefore more aggressive treatment of the estimated tumor extent might potentially improve patient survival. Further studies are needed to elucidate the ultimate value of these estimated model parameters.

Deformable registration of the atlas to the patient scans, that is jointly estimated with segmentation, allowed us to construct a statistical atlas of the glioma. To that end, we mapped the tumor segmentations of patients scans into the common space by inverting the estimated deformation fields. We show that the average tumor probability is a nonuniform function which reaches a maximum in the left temporal lobe within our patient population. Such atlases can potentially carry prognostic value, as gliomas are believed to progress preferentially along white matter fiber pathways as well as along brain vasculature. Such relationships can be investigated in a principled way in the future using our atlas.

In summary, application of GLISTR to a large number of data sets reveals that the method can produce promising segmentation and atlas registration results, even in the presence of large mass-effects, necrosis and edema. GLISTR is freely available for download through the homepage of Section of Biomedical Image Analysis, the University of Pennsylvania.[3]

Acknowledgment

The authors would like to thank A. Schuh for his excellent efforts on packaging and improving the quality of the software.

This work was supported by the National Institutes of Health (NIH) under Grant R01 NS042645.

Appendix

We now derive the update equation for the deformation field **h** as described in (19). Keeping the current estimation of $\mathbf{q}'$ and $\mathbf{\Phi}'$ fixed, we compute the variation of (11) with respect to **h** in the direction of infinitely small and arbitrary test function **v** and assign it to zero [59]. i.e.,

$$\langle \partial Q(\mathbf{\Phi}', \mathrm{h}, \mathbf{q}' | \mathbf{\Phi}', \mathrm{h}', \mathbf{q}'), \mathbf{v} \rangle = \lim_{\|\mathbf{v}\| \to 0} \{ Q(\mathbf{\Phi}', \mathrm{h}' + \mathbf{v}, \mathbf{q}' | \mathbf{\Phi}', \mathrm{h}', \mathbf{q}') - Q(\mathbf{\Phi}', \mathrm{h}', \mathbf{q}' | \mathbf{\Phi}', \mathrm{h}', \mathbf{q}') \} = 0. \quad (22)$$

where $\mathbf{h}' + \mathbf{v} \equiv \{\mathbf{h}'(\mathbf{x}) + \mathbf{v}(\mathbf{x}) | \mathbf{x} \in \Omega_F\}$. We replace (11) in (22), and consider only those terms that depend on deformation parameters $\mathbf{h}'(\mathbf{x})$. Abbreviating $\pi_k(\mathbf{h}'(\mathbf{x})) = \pi_k \circ \mathbf{h}'$, and $\mathbf{v} = \mathbf{v}(\mathbf{x})$, (22) can be written as

[3]Available online: http://www.rad.upenn.edu/sbia/projects/glistr.html

$$\langle \partial Q(\mathbf{\Phi}', \mathbf{h}, \mathbf{q}' | \mathbf{\Phi}', \mathbf{h}', \mathbf{q}'), \mathbf{v} \rangle = \sum_{\mathbf{x} \in \Omega_F} \sum_{k=1}^{K} p_k(\mathbf{x}) \left[\log \left(\pi_k \circ \mathbf{h}' + \mathbf{v}^t \cdot \nabla \pi_k \circ \mathbf{h}' + \frac{1}{2} \mathbf{v}^t \mathbf{H}(\pi_k \circ \mathbf{h}') \mathbf{v} \right) - \log(\pi_k \circ \mathbf{h}') \right] \quad (23)$$

where ∇ and $\mathbf{H}$ stand for the gradient and Hessian operators. Considering the fact that

$$\log(x+\varepsilon) = \log(x) + \frac{\varepsilon}{x} - \frac{1}{2} \frac{(\varepsilon)^2}{x^2}, \|\varepsilon\| \to 0$$

and keeping the terms up to second order with respect to $|\mathbf{v}|$ we can approximate the first logarithm term using

$$\log \left(\pi_k \circ \mathbf{h}' + \mathbf{v}^t \cdot \nabla \pi_k \circ \mathbf{h}' + \frac{1}{2} \mathbf{v}^t \mathbf{H}(\pi_k \circ \mathbf{h}') \mathbf{v} \right) = \log(\pi_k \circ \mathbf{h}') + \frac{\mathbf{v}^t \cdot \nabla \pi_k \circ \mathbf{h}' + \frac{1}{2} \mathbf{v}^t \mathbf{H}(\pi_k \circ \mathbf{h}') \mathbf{v}}{\pi_k \circ \mathbf{h}'} - \frac{\mathbf{v}^t \cdot \nabla \pi_k \circ \mathbf{h}' (\nabla \pi_k \circ \mathbf{h}')^t \cdot \mathbf{v}}{2(\pi_k \circ \mathbf{h}')^2} \quad (24)$$

plugging (24) back into (23), we arrive at

$$\langle \partial Q(\mathbf{\Phi}', \mathbf{h}, \mathbf{q}' | \mathbf{\Phi}', \mathbf{h}', \mathbf{q}'), \mathbf{v} \rangle = \sum_{\mathbf{x} \in \Omega_F} \mathbf{v}^t . \{ \mathbf{r} + \mathbf{W} . \mathbf{v} \} = 0 \quad (25)$$

where $\mathbf{W}$ and $\mathbf{r}$, have the same definitions in equations (16) and (17). Since (25) should hold for every $\mathbf{v}$, we must have: $\mathbf{r} + \mathbf{W}.\mathbf{v} = 0$, from which (18) the update equation of deformation field in (19) can be justified.

REFERENCES

1. Collins V. Gliomas. Cancer Survey. 1998; 32:37–51.
2. Markert, J.; Devita, V.; Rosenberg, S.; Hellman, S. Gliobalastoma Multiforme. Burlington, MA: Jones Bartlett; 2005.
3. Weizman L, Sira L, Joskowicz L, Constantini S, Precel R, Shortly B, Bashat D. Automatic segmentation, internal classification, and follow-up of optic pathway gliomas in MRI. Med. Image Anal. 2012; 16:177–188. [PubMed: 21852179]
4. Liu J, Udupa J, Odhner D, Hackney D, Moonis G. A system for brain tumor volume estimation via MR imaging and fuzzy connectedness. Comput. Med. Imag. Graphics. 2005; 29:21–34.
5. Henson J, Ulmer S, Harris G. Brain tumor imaging in clinical trials. Am. J. Neuroradiol. 2008; 29:419–424. [PubMed: 18272557]
6. Phillips W, Velthuizen R, Phupanich S, Hall L, Clarke L, Silbiger M. Application of fuzzy C-means segmentation technique for tissue differentiation in MR images of a hemorrhagic glioblastoma multiforme. Magn. Reson. Imag. 1995; 29:277–290.
7. Clark M, Hall L, Goldgof D, Velthuizen R, Murtagh R, Silbiger M. Automatic tumor segmentation using knowledge-based techniques. IEEE Trans. Med. Imag. 1998; 17:187–201.
8. Fletcher-Heath L, Hall L, Goldgof D, Murtagh F. Automatic segmentation of nonenhancing brain tumors in magnetic. Artif. Intell. Med. 2001; 21:43–63. [PubMed: 11154873]
9. Lefhon A, Cates J, Whitaker R. GPU-based level sets for 3-D brain tumor segmentation. Proceeding of Medical Image Computig and Computer Aided Intervation (MICCAI). 2003:568–572.
10. Ho S, Bullit E, Gerig G. Level-set evolution with region competition: Automatic 3-D segmentation of brain tumors. Proc. Int. Conf. Pattern Recognit. 2002:532–535.
11. Cobzas D, Birkbeck N, Schmidt M, Jagersand M. 3-D variational brain tumor segmentation using a high dimensional feature set. Proc. IEEE Int. Conf. Comput. Vis. (ICCV). 2007:1–8.

12. Lee C, Wang S, Murtha A, Greiner R. Segmenting brain tumors using pseudo conditional random fields. Proc. Med. Image Comput. Computer Aided Intervent. (MICCAI). 2008:359–366.
13. Zhang J, Ma K, Er M, Chong V. Tumor segmentation from magnetic resonance imaging by learning via one-class support vector machine. Int. Workshop Adv. Image Technol. 2004:207–211.
14. Verma R, Zacharaki E, Ou Y, Cai H, Chawla S, Lee S, Melhem E, Wolf R, Davatzikos C. Multiparametric tissue characterization of brain neoplasms and their recurrence using pattern classification of MR images. Acad. Radiol. 2008; 15:566–977.
15. Ayachi R, Amor N. Brain tumor segmentation using support vector machines. Proc. Eur. Conf. Symbolic Quantitative Approach. Reasoning Uncertainty. 2009:736–747.
16. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm. IEEE Trans. Med. Imag. 2001; 20:45–57.
17. Prastawa M, Bullitt E, Gerig G. A brain tumor segmentation framework based on outlier detection. Med. Image Anal. 2004; 8(3):275–283. [PubMed: 15450222]
18. Wels M, Carneiro G, Aplas A, Huber M, Hornegger J, Comaniciu D. A discriminative model-constrained graph-cuts approach to fully automated pediatric brain tumor segmentation in 3-D MRI. Proc. Med. Image Comput. Computer Aided Intervent. (MICCAI). 2008:67–75.
19. Gorlitz, L.; Menze, B.; Weber, M.; Kelm, B.; Hamprecht, F. Semi-Supervised Tumor Detection in Magnetic Resonance Spectroscopic Images using Discriminative Random Fields. Hamprecht, FA.; Schnorr, C.; Jahne, B., editors. New York: Springer; 2007. p. 224-233.Lecture Notes in Computer Sciences
20. Moon N, Bullit E, Leemput K, Gerig G. Model-based brain and tumor segmentation. Proc. Int. Conf. Pattern Recognit. 2002:528–531.
21. Gering D, Grimson W, Kikinis R. Recognizing deviation from normalcy for brain tumor segmentation. Proc. Med. Image Comput. Computer Aided Intervent. (MICCAI). 2002:388–395.
22. Prastawa M, Bullitt E, Moon N, Leemput K, Gerig G. Automatic brain tumor segmentation by subject specific modification of atlas priors. Acad. Radiol. 2003; 10:1341–1348. [PubMed: 14697002]
23. Corso J, Sharon E, Dube S, Suden S, Sinha U, Yuille A. Efficient multilevel brain tumor segmentation with integrated Bayesian model classification. IEEE Trans. Med. Imag. 2008 May; 27(5):629–731.
24. Menze B, Leemput K, Lashkari D, Weber M, Ayache N, Golland P. A generative model for brain tumor segmentation in multi-modal images. Proceeding of Medical Image Computing and Computer Aided Intervention (MICCAI). 2010:151–159.
25. Duda, R. Pattern Classification. New York: Wiley; 2001.
26. Tu Z. Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering. IEEE Trans. Med. Imag. 2005:1589–1596.
27. Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Trans. Med. Imag. 2004 Sep.26(9):1124–1137.
28. Sharon E, Galun M, Sharon D, Basri R, Brandt A. Hierarchy and adaptivity in segmenting visual scenes. Nature. 2006; 442(7104):810–813. [PubMed: 16810176]
29. Kindermann, R.; Snell, J. Markov Random fields and their Applications. Providence, RI: Am. Math. Soc.; 1980.
30. Pohl K, Fisher J, Kikinis R, Wells W. A Bayesian model for joint segmentation and registration. Neuroimage. 2006; 31:228–239. [PubMed: 16466677]
31. Ashburner J, Friston K. Unified segmentation. NeuroImage. 2005; 26:839–851. [PubMed: 15955494]
32. Leemput K, Maes F, Vandermeulen D, Colchester A, Suetens P. Automated segmentation of multiple sclerosis lesions by model outlier detection. IEEE Trans. Med. Imag. 2001 Aug.20(8): 677–688.
33. Hogea C, Davatzikos C, Biros G. An image-driven parameter estimation problem for a reaction-diffusion glioma growth model with mass effects. J. Math. Biol. 2008; 56:793–825. [PubMed: 18026731]

34. Swanson K, Alvord E, Murray D. A quantitative model for differential motality of gliomas in grey and white matter. Cell Proliferation. 2000; 33:317–329. [PubMed: 11063134]
35. Gooya A, Biros G, Davatzikos C. Deformable registration of glioma images using EM algorithm and diffusion reaction modeling. IEEE Trans. Med. Imag. 2011 Feb.30(2):375–390.
36. Mohamed A, Zacharaki E, Shen D, Davatzikos C. Deformable registration of brain tumor images via a statistical model of tumor-induced deformation. Med. Image Anal. 2006; 10:752–763. [PubMed: 16860588]
37. Zacharaki E, Hogea C, Shen D, Biros G, Davatzikos C. Non-diffeomorphic registration of brain tumor images by simulating tissue loss and tumor growth. Neuroimage. 2009; 46:762–774. [PubMed: 19408350]
38. Gooya A, Pohl K, Billelo M, Biros G, Davatzioks C. Joint segmentation and deformable registration of brain scans guided by a tumor growth model. Proceeding of Medical Image Computig and Computer Aided Intervention (MICCAI). 2011:532–540.
39. Hogea, C.; Abraham, F.; Bios, G.; Davatzikos, C. Medical Image Computing and Computer Aided Intervention (MICCAI) Workshop on Biomechanics. Copenhagen, Denmark: 2006. A framework for soft tissue simulations with applications to modeling brain tumor mass-effect in 3-D images.
40. Clatz O, Sermesant M, Bondiau P, Delingette H, Warfield S, Mal G, Ayache N. Realistic simulation of the 3-D growth of brain tumors in MR images coupling diffusion with mass effect. IEEE Trans. Med. Imag. 2005 Dec.24(10):1334–1346.
41. Konukoglu E, Clatz O, Bondiau P, Delingette H, Ayache N. Extrapolating glioma invasion margin in brain magnetic resonance images: Suggesting new irradiation margins. Med. Image Anal. 2010; 14(2):111–125. [PubMed: 20042359]
42. Konukoglu E, Clatz O, Menze B, Stieltjes B, Weber M, Mandonnet E, Delingette H, Ayache N. Image guided personalization of reaction-diffusion type tumor growth models using modified anisotropic eikonal equations. IEEE Trans. Med. Imag. 2010 Jan.29(1):77–95.
43. Landau, L.; Lifshitz, E. Theory of Elasticity. Oxford, UK: Butterworth Heinemann; 1986.
44. Hogea C, Davatzikos C, Biros G. Brain-tumor interaction biophysical models for medical image registration. SIAM J. Sci. Comput. 2008; 30(6):3050–3072.
45. Hagemann A, Rohr K, Stiehl H, Spetzger U, Gilabach J. Biomechanical modeling of the human head for physically based, nonrigid image registration. IEEE Trans. Med. Imag. 1999 Oct.18(10): 875–884.
46. Dempster A, Laird N, Rubin D. Maximum likelihood from incomplete data via the EM algorithm. JR. Stat. Soc. Series B. 1977; 39(1):1–38.
47. Bishop, C. Pattern Recognition and Machine Learning. New York: Springer; 2009.
48. Thirion J. Image matching as diffusion process: An analogy with Maxwell's demons. Med. Image Anal. 1998; 2:243–260. [PubMed: 9873902]
49. Hough P, Kolda T, Torczon V. Asynchronous parallel pattern search for nonlinear optimization. SIAM J. Sci. Comput. 2001; 23(1):134–156.
50. Sled J, Zijdenbos A, Evans A. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans. Med. Imag. 1998 Feb.17(1):81–97.
51. Smith S. Fast robust automated brain extraction. Human Brain Mapp. 2002; 17:143–155.
52. Jenkinson M, Bannister P, Brady J, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage. 2002; 17:825–841. [PubMed: 12377157]
53. Chang C, Lin C. LIBSVM: A library for support vector machines 2001 [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm.
54. Dice L. Measures of the amount of ecologic association between species. Ecology. 1945; 26(3): 297–302.
55. Bouix, S.; Ungar, L.; Dickey, C.; McCarley, RW.; Shenton, M. Proc. Med. Image Comput. Computer Aided Intervent. (MICCAI). Saint Malo, France: 2004. Evaluating automatic brain tissue classifiers; p. 1038-1039.
56. Conn, MP. Neuroscience in Medicine. New York: Humana Press; 2008.

57. Davatzikos C, Genca A, Xua D, Resnickb S. Voxel-based morphometry using the ravens maps: Methods and validation using simulated longitudinal atrophy. NeuroImage. 2001; 14:1361–1369. [PubMed: 11707092]

58. Jerome, M.; Well, A. Research Design and Statistical Analysis. Mahwah, NJ: Lawrence Erlbaum; 2003.

59. Hildebrand, F. Advanced Calculus for Applications. Englewood Cliffs, NJ: Prentice-Hall; 1976.

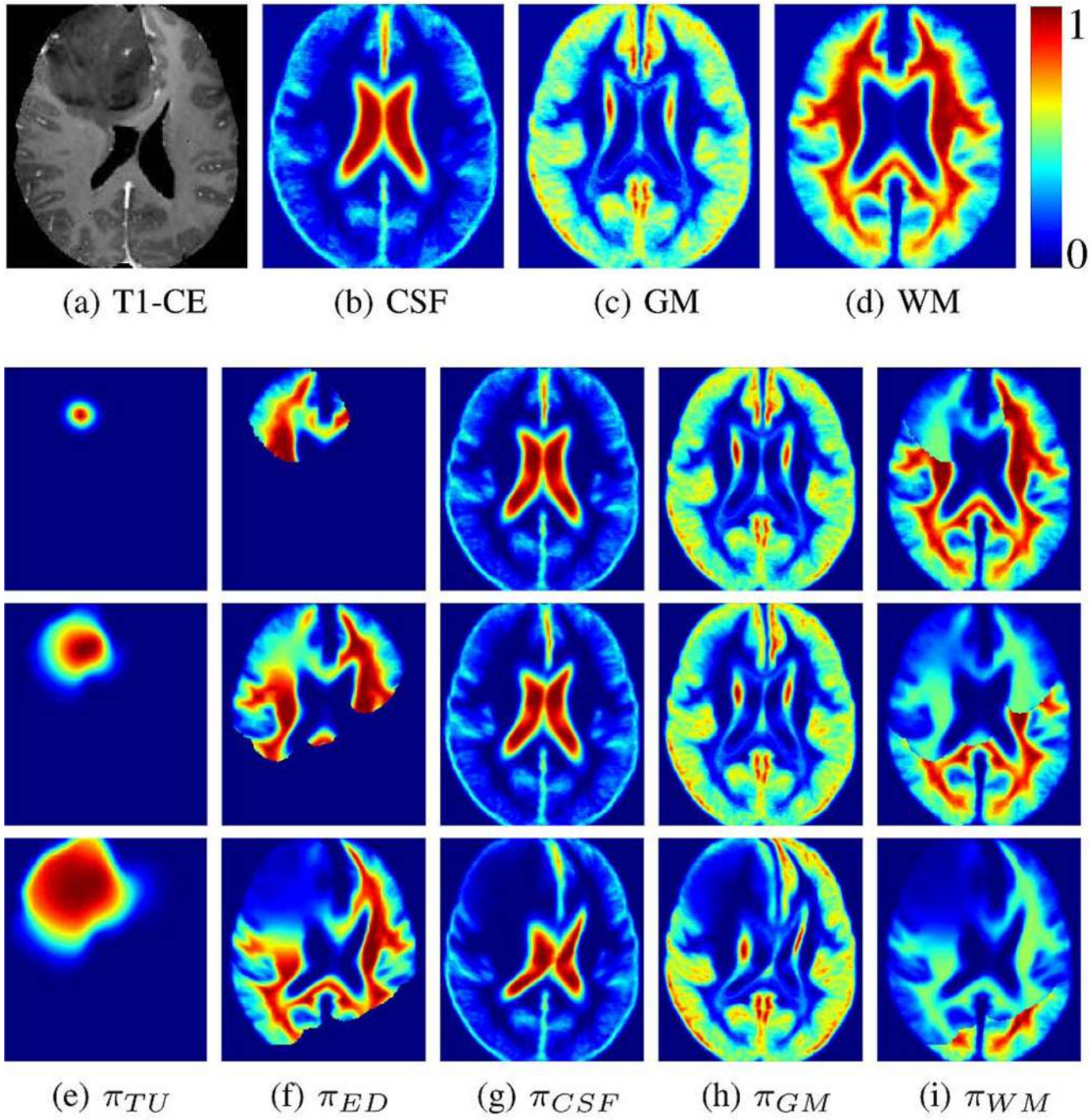

Fig. 1.

(a) Sample glioma scan, (b)–(d) healthy cerebro spinal fluid, gray matter and white matter probability maps, evolution of the corresponding probability maps of (e) tumor, (f) edema, (g) cerebro spinal fluid, and (h) gray matter (i) white matter at $t \simeq 0$ (right after initial condition), $t = T/2$, and $t = T$, computed by the proposed method.

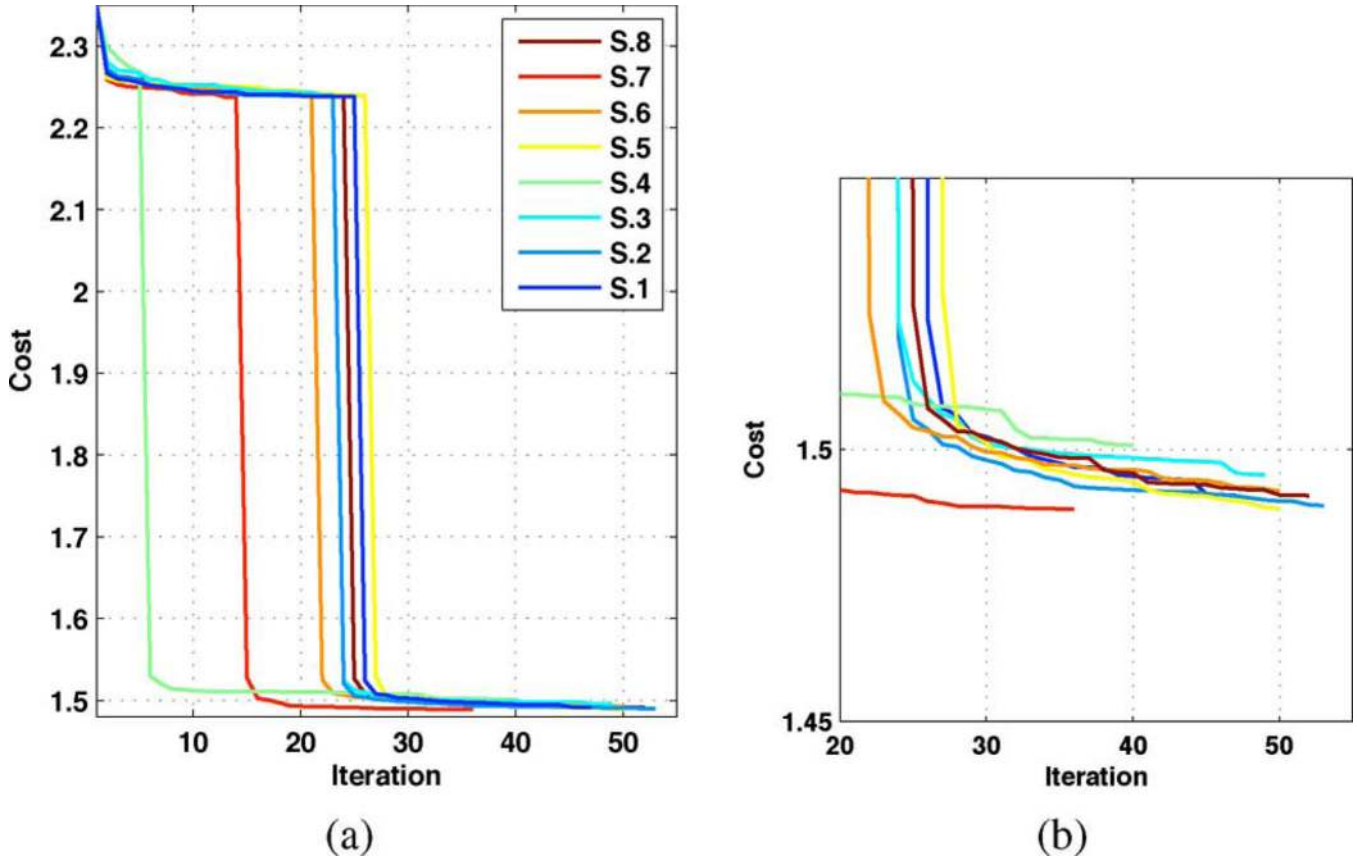

Fig. 2.

(a) Minimization of the total cost (defined as the negative of Q in the M-Step) for eight different initial seed points. Initial variations of the cost correspond to updates in $\mathbf{q}$ (with $\mathbf{h}$ and $\mathbf{\Phi}$ fixed) whereas the large drops correspond to first round of convergence and update in $\mathbf{h}$ (with $\mathbf{q}$ fixed), (b) enlargement of the second and third rounds of optimization with minor changes of the cost, showing the convergence.

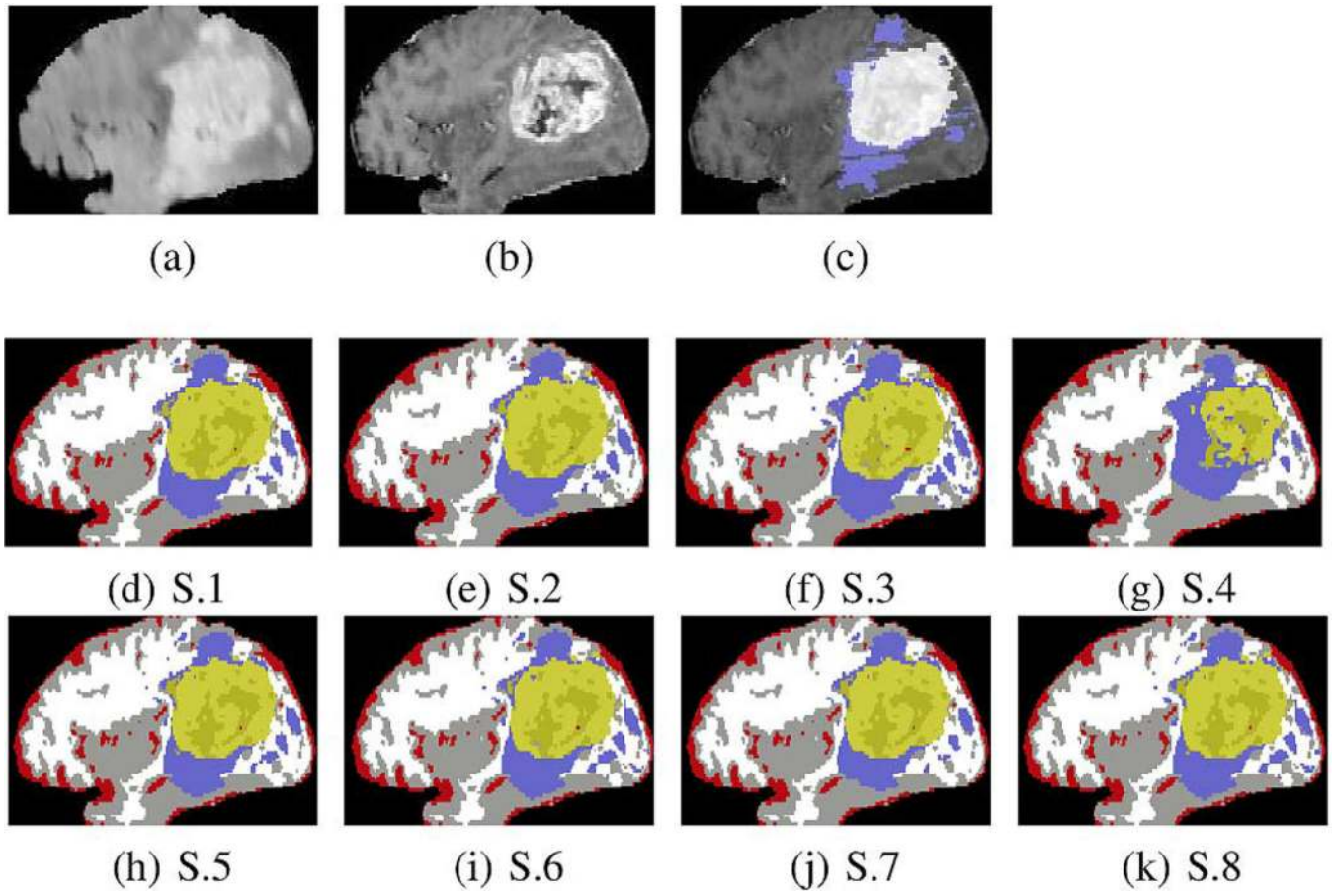

Fig. 3.

(a)–(b) Slices of FLAIR and T2 modalities, (c) the corresponding slice of the reference segmentations from our clinician expert, (d)–(k) the corresponding final segmentations acquired with GLISTR for eight different initial seed points (S.1–S.8 in Fig. 2). Except for S.4, all other segmentations look reasonably similar to the references in (c). This shows that the method is relatively robust to changes in initial seed location.

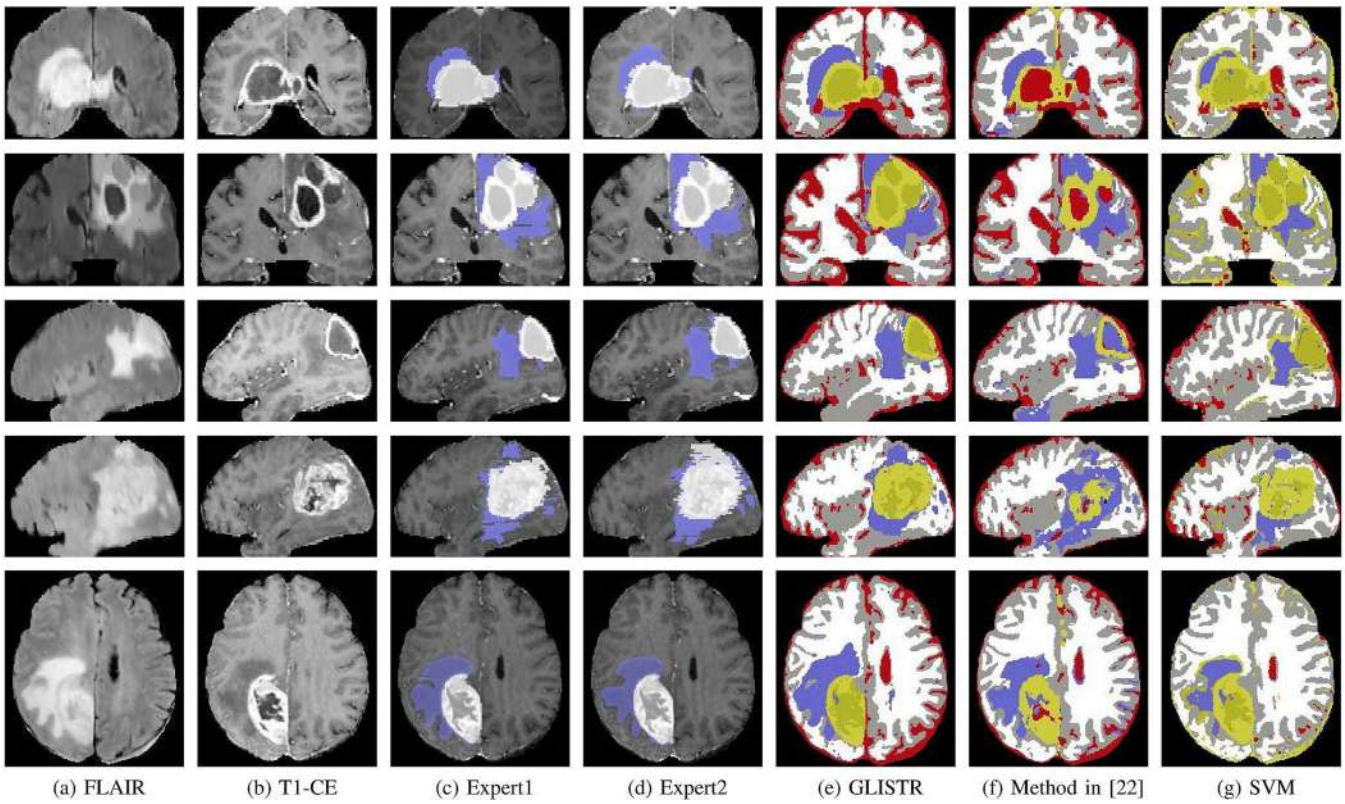

Fig. 4.

Comparison of segmentations obtained from various raters. Each row corresponds to a single patient: FLAIR and T1-CE modalities are shown in columns (a) and (b), followed by segmentations of Expert1, Expert2, the proposed method (GLISTR), methods proposed in [22] and [14] (SVM) in columns (c)–(g) respectively. Comparison of (c) to (d) reveals the inter-rater variability of the segmentations (e.g., compare second row from bottom; a larger tumor volume has been delineated by Expert2). Many necrotic regions are not segmented as tumor by [22], whereas a lot of false-tumor positives are observable in SVM results. The proposed GLISTR method, however, has the most resemblance to the results of Expert1 considered as the reference volumes in this study.

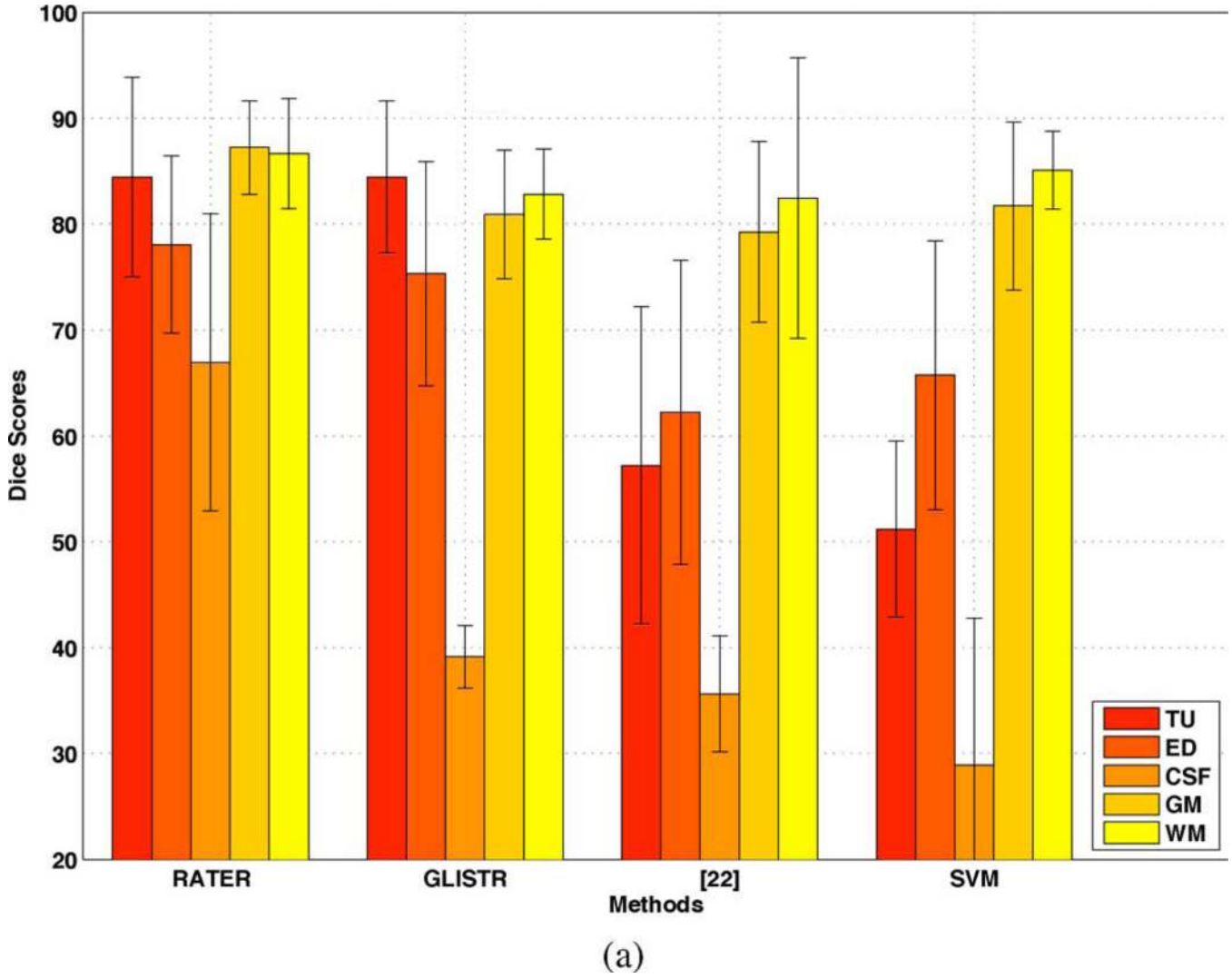

Fig. 5.

Error bar graphs of the measured Dice scores for the segmented labels for various raters, the second human rater (RATER), the proposed method (GLISTR), method of [22], and SVM based method [14], with regard to the first rater as the ground truth. On tumor and edema labels GLISTR outperforms these methods and reaches a similar accuracy of the second clinical expert.

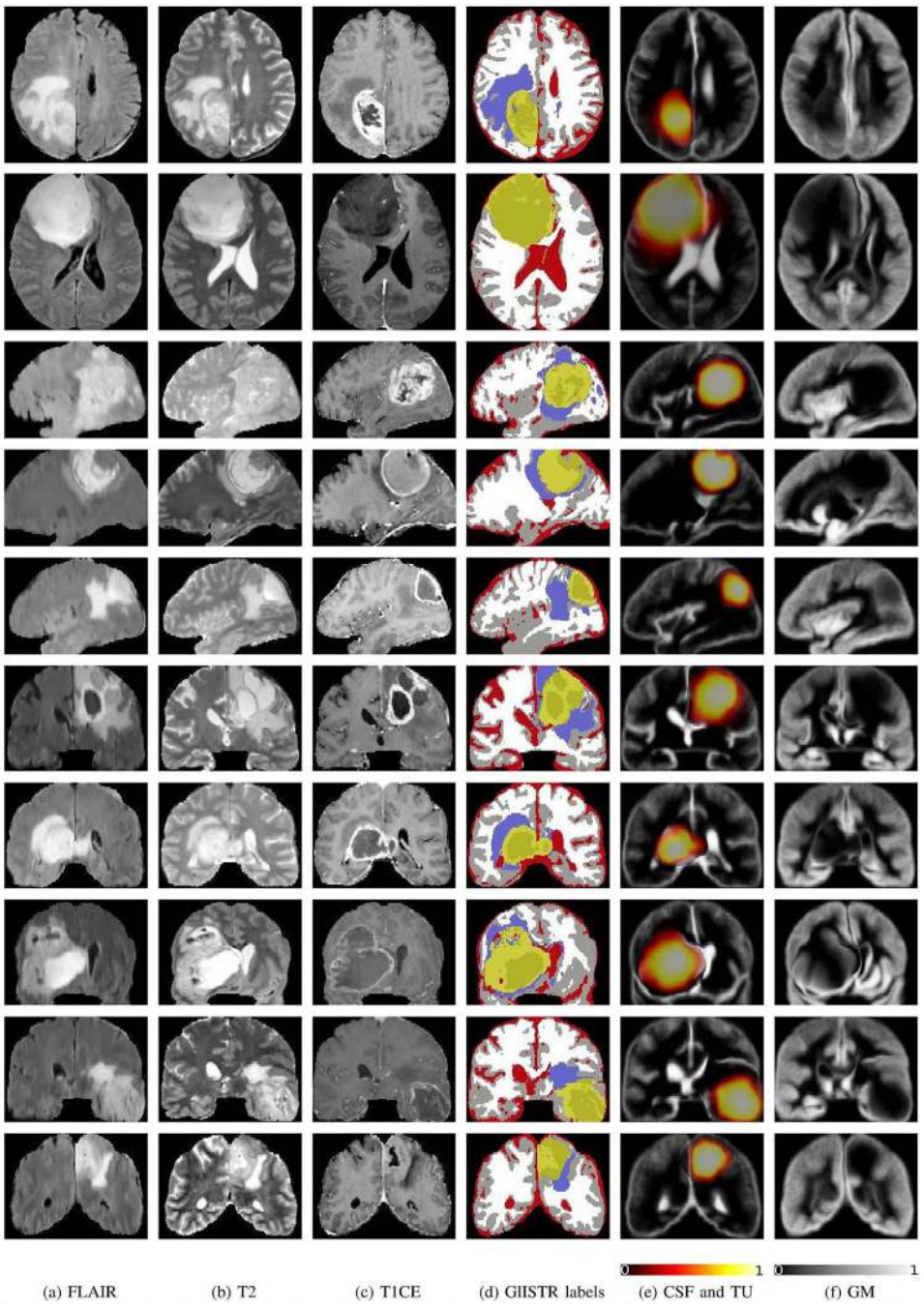

Fig. 6.

Segmentation and registration results for 10 sample patients. Each row corresponds to a single patient and represents the results in the slice with largest tumor section. (a)–(c) FLAIR, T2 and T1-CE images, (d) segmentation results of GLISTR showing enhancing tumor, necrosis, edema, CSF, gray and white matters in light and dark yellows, purple, red, gray and white colors respectively, (e) overlay of the tumor and CSF probability maps registered to the patient scans, (f) probability map of GM registered to the patient scans.

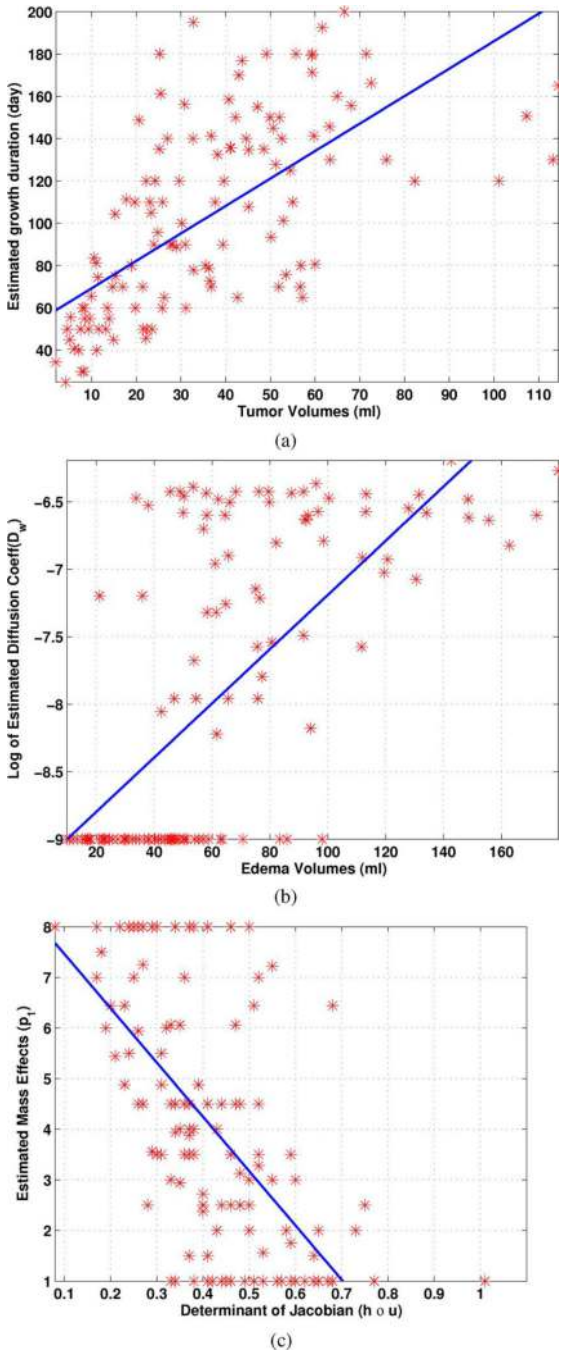

Fig. 7.

Scatter plots of the estimated tumor model parameters for the 122 glioma cases. (a) Tumor growth duration T versus the volume of the segmented tumor $Vol(TU)$, (b) Diffusion coefficient D_w versus the volume of the segmented edema $Vol(ED)$, (c) The average determinant of the Jacobian of the total composite deformation field $|Jac(\mathbf{h} \circ \mathbf{u})|$ in the tumor area versus mass-effect coefficient p_1 The blue line fits show the linearities in the relations of the estimated parameters. Further deviations from these lines correspond to higher

degrees of the nonlinearity. Hence, the most linear relation is observed between T and $Vol(TU)$, whereas the least linearity is observed for D_w versus $Vol(ED)$ (see Section IV-E).

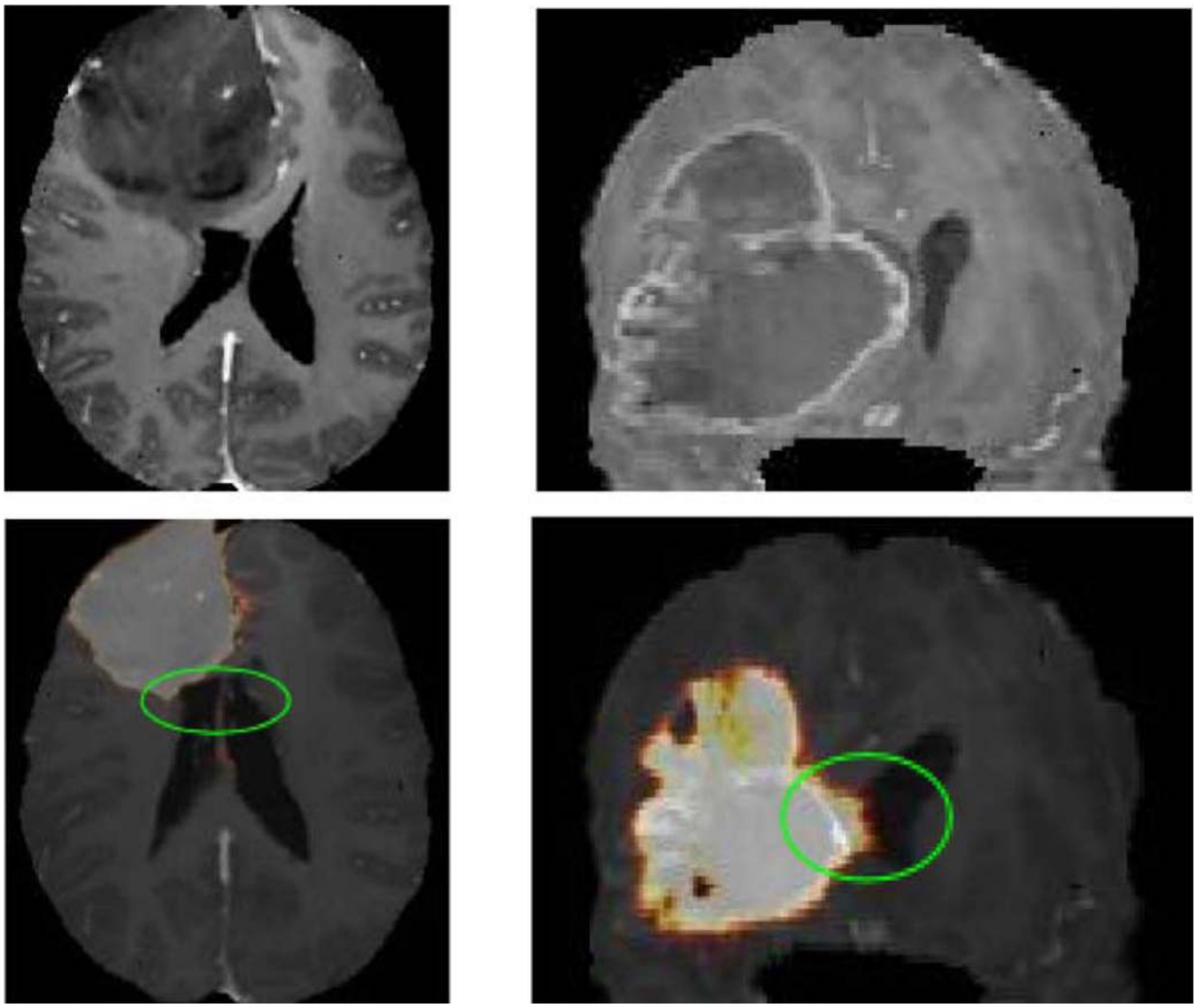

Fig. 8.

First row: sample T1CE images of gliomas, Second row: atlas normalized tumor posteriors overlayed on their corresponding T1CE images. Ventricles seem to be relaxed compared to the original patients scans (*cf.* the regions encircled by the markers).

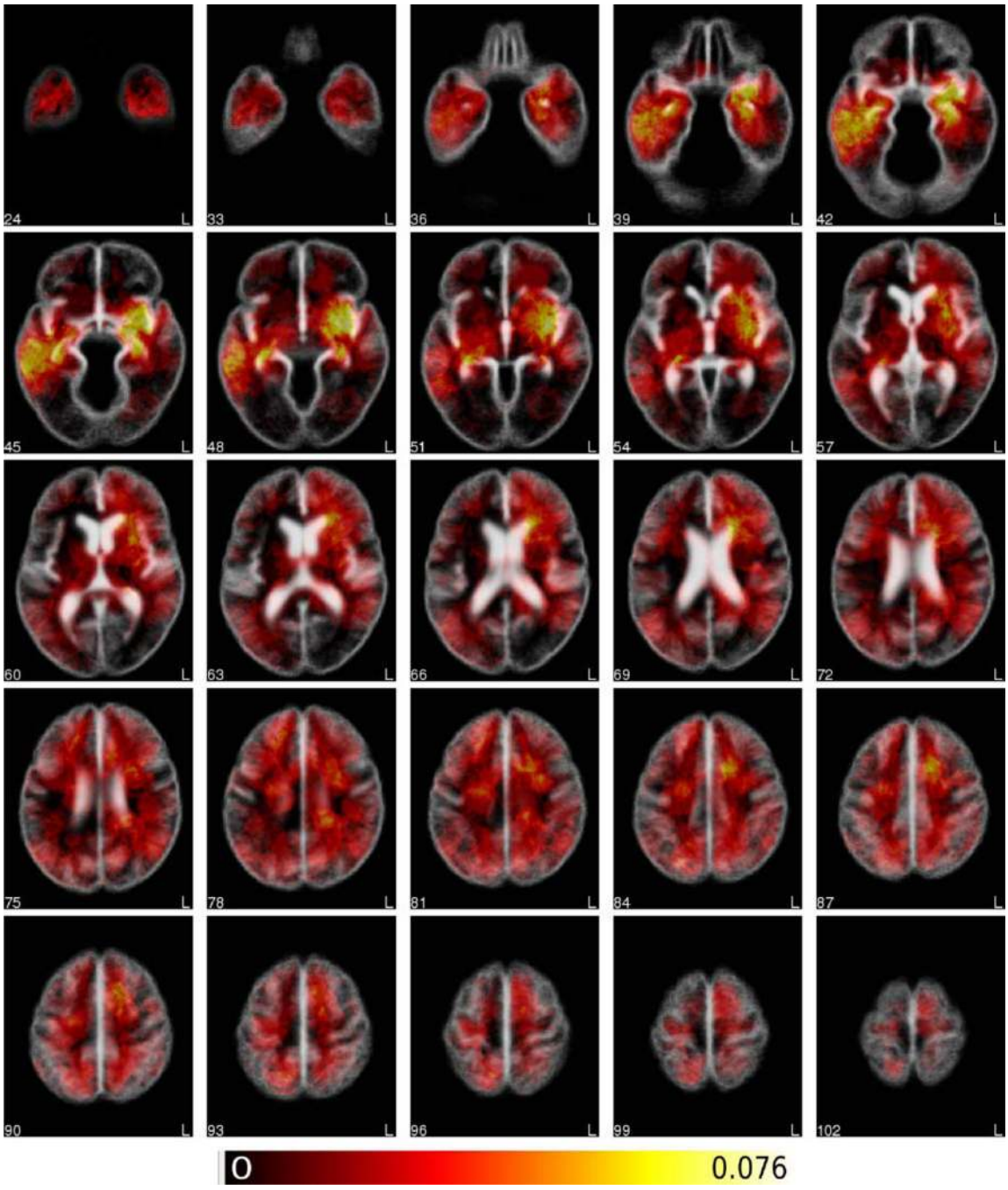

Fig. 9.

Tumor atlas indicating the spatial probability distribution of tumor on the atlas space. We mapped the posterior probability of tumors back to the atlas space by inverting the total deformation field $(\mathbf{h} \circ \mathbf{u})^{-1}$. Then we computed the average of these transformed tumor posteriors over 122 glioma cases. The color bar indicates that, within our patient population, the region with the highest tumor probability is placed in the left temporal lobe of the brain.

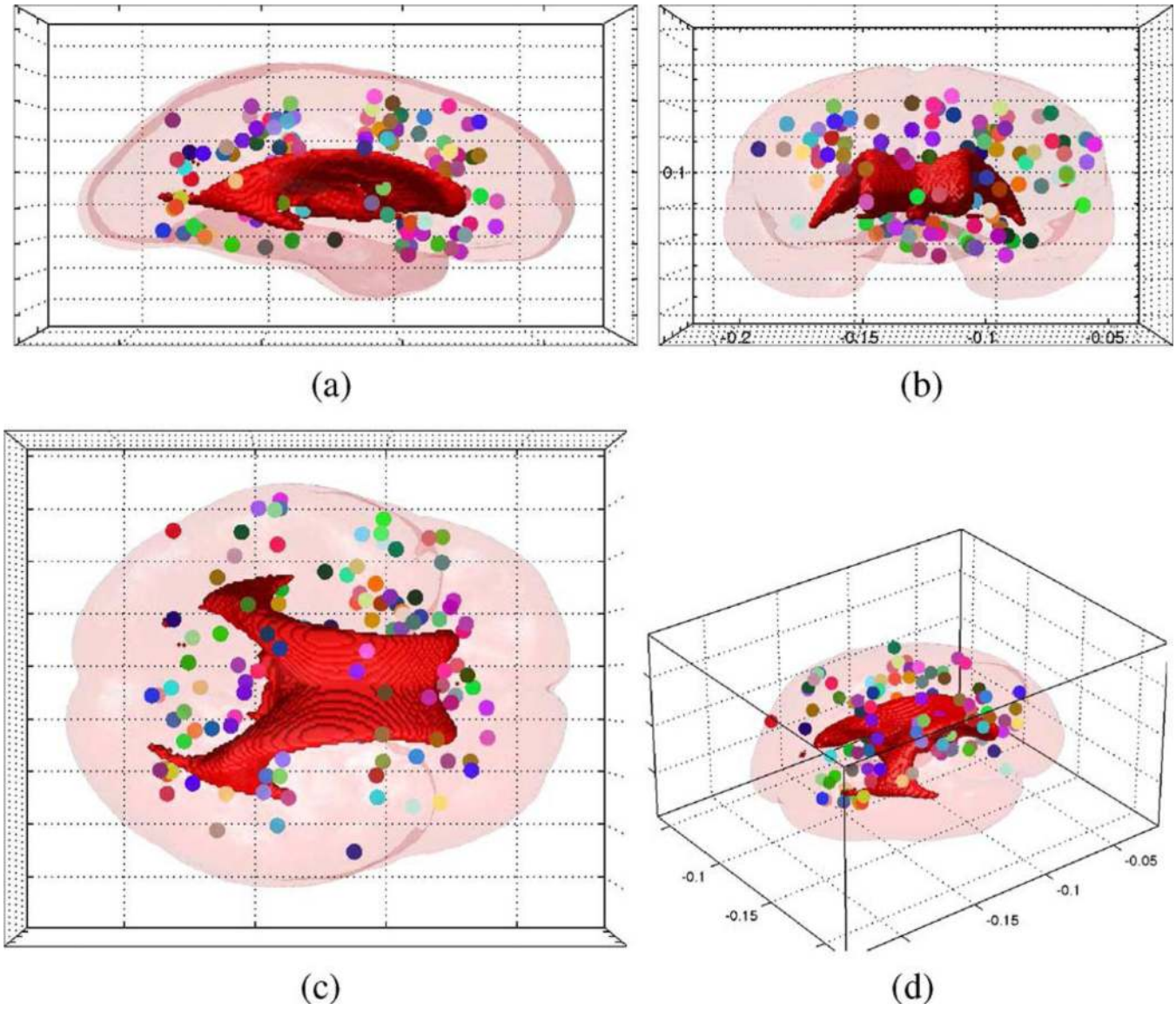

Fig. 10.

Spatial distribution of the estimated tumor seeds for 122 glioma cases in the atlas space indicated from: (a) sagital, (b) coronal, (c) axial, and (d) 3-D views. The colors distinguish the spatial location of each tumor seed in different views. Fewer seeds are located in the frontal and occipital lobes compared to other regions of the brain.

TABLE I

Converged Cost Values And Corresponding Dice Overlap Ratios for Eight Different Seed Locations

Seed	Converged Cost	TU Dice(%)	ED Dice(%)
S.1	1.4960	87.40	60.15
S.2	1.4897	87.40	59.76
S.3	1.4953	88.11	57.76
S.4	1.5008	73.14	54.61
S.5	1.4892	86.88	58.94
S.6	1.4924	87.87	58.53
S.7	1.4890	87.90	59.46
S.8	1.4916	87.06	57.57

TABLE II

Surface Distances (*MM*) Between the Reference and the Acquired Segmented Volumes by Different Methods for Tumor (TU) and (ED) Labels

	RATER	GLISTR	Method proposed in [22]	SVM
TU	2.78±1.80	2.00±0.29	9.35±7.35	26.45±8.44
ED	2.45±1.31	2.28±1.29	7.7±7.00	2.56±1.73

TABLE III

Spearman Correlation Coefficients and p-Values for: T Versus the Volume of the Segmented Humor $Vol(TU)$, Diffusion Coefficient D_w Versus the Volume of the Segmented Edema $Vol(ED)$ the Average Determinant of the Jacobian of the Total Composite Deformation Field in The Tumor Area $|Jac(\mathbf{h} \circ \mathbf{u})|$ Versus Mass-Effect Coefficient p_1, $|Jac(\mathbf{u})|$ and $|Jac(\mathbf{h})|$

	Correlation coefficient	p-value
T, $Vol(TU)$	0.7272	$1.6773e^{-21}$
D_w, $Vol(ED)$	0.6616	$1.1102e^{-16}$
p_1, $\|Jac(\mathbf{h} \circ \mathbf{u})\|$	−0.6383	$2.5535e^{-15}$
$\|Jac(\mathbf{u})\|$, $\|Jac(\mathbf{h} \circ \mathbf{u})\|$	0.6104	$8.3711e^{-14}$
$\|Jac(\mathbf{h})\|$, $\|Jac(\mathbf{h} \circ \mathbf{u})\|$	0.4701	$4.4820e^{-08}$