PORTR: Pre-Operative and Post-Recurrence Brain Tumor Registration

Dongjin Kwon*, *Member, IEEE,* Marc Niethammer, *Member, IEEE,* Hamed Akbari, Michel Bilello, Christos Davatzikos, *Fellow, IEEE,* and Kilian M. Pohl, *Member, IEEE*

***Abstract*—We propose a new method for deformable registration of pre-operative and post-recurrence brain MR scans of glioma patients. Performing this type of intra-subject registration is challenging as tumor, resection, recurrence, and edema cause large deformations, missing correspondences, and inconsistent intensity profiles between the scans. To address this challenging task, our method, called *PORTR*, explicitly accounts for pathological information. It segments tumor, resection cavity, and recurrence based on models specific to each scan. *PORTR* then uses the resulting maps to exclude pathological regions from the image-based correspondence term while simultaneously measuring the overlap between the aligned tumor and resection cavity. Embedded into a symmetric registration framework, we determine the optimal solution by taking advantage of both discrete and continuous search methods. We apply our method to scans of 24 glioma patients. Both quantitative and qualitative analysis of the results clearly show that our method is superior to other state-of-the-art approaches.**

***Index Terms*—Brain tumor MRI, deformable registration, tumor segmentation, tumor growth model, discrete-continuous optimization**

I. Introduction

THE treatment of brain gliomas could greatly benefit from discovering imaging markers in the pre-operative scans that accurately predict tumor infiltration and subsequent tumor recurrence [1]–[3]. One possible approach for discovering these markers is to first align the pre-operative and the post-recurrence structural brain MR scans of a patient and then to analyze the imaging characteristics of tissue that later turn into tumor recurrence [4], [5]. This strategy relies on accurate registration as the size of tumor recurrence is usually not very large. However, non-rigid registration of the pre-operative and post-recurrence scans is very challenging due to large deformations, missing correspondences, and inconsistent intensity profiles between the scans. The large deformations and missing correspondences are due to the glioma in the pre-operative scans causing large mass effects [6] as well as the resection cavities and tumor recurrence in the post-recurrence scans, which are acquired several months or years after surgery. The inconsistent intensity profiles result from tissue labeled as edema in the pre-operative scan transforming to healthy tissue in the post-recurrence scan (and vice versa). Thus, corresponding regions can have very different intensity profiles. Fig. 1 shows a typical case with the anatomy around the tumor being confounded by resection cavity, tumor recurrences, and edema. In this paper, we develop a registration method to cope with the missing correspondence issue between the scans and show that the results are much more accurate than general-purpose registration methods.

Existing registration methods mostly deal with the missing correspondence issue by excluding the pathology during the mapping process [7]–[9]. They require segmentations of the brain scans, such as Clatz *et al.* [10], who align the pre-operative and intra-operative brain scans by first matching selected regions of healthy tissue. The approach then applies a bio-mechanical model to the resulting map to interpolate the deformations to the remaining image domain. Risholm *et al.* [11] avoid the prerequisite of a label map by alternating between extracting the resection area and estimating deformations. However, it is difficult to estimate gross deformation on brain glioma scans by excluding the pathology. To take explicit account of the tumor region, approaches register tumor scans to a healthy brain template by simulating mass effects of the tumor on the template [12], [13]. Alternatively, generative models, such as Prastawa *et al.* [14] and Menze *et al.* [15], inject a prior of the tumor into an atlas of a healthy population and segment the tumor by aligning this atlas to the scan. Gooya *et al.* [16] extend this idea to non-rigid registration by growing the tumor inside the atlas until the deformed atlas resembles the pathology and healthy tissue shown in the brain tumor scan.

To the best of our knowledge, this paper proposes the first approach specifically targeted towards the registration of pre-operative and post-recurrence glioma scans, called *Pre-Operative and post-Recurrence brain Tumor Registration* (*PORTR*). One could register an atlas to each scan individually, for example via [16], and then concatenate the corresponding registrations [17]. However, this approach ignores the fact that the scans are from the same patient. It thus has to solve the

Manuscript received September 21, 2013; revised November 23, 2013; accepted November 25, 2013. This work was supported by the National Institutes of Health (NIH) under Grant R01 NS042645. *Asterisk indicates corresponding author.*

Copyright (c) 2013 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

*D. Kwon is with the Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA. (e-mail: Dongjin.Kwon@uphs.upenn.edu)

M. Niethammer is with the Department of Computer Science and Biomedical Research Imaging Center, School of Medicine, University of North Carolina, Chapel Hill, NC, USA.

H. Akbari, M. Bilello, and C. Davatzikos are with the Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA.

K.M. Pohl is with the Department of Psychiatry and Behavioral Sciences, Stanford University, and the Center for Health Sciences, SRI International, Menlo Park, CA, USA.

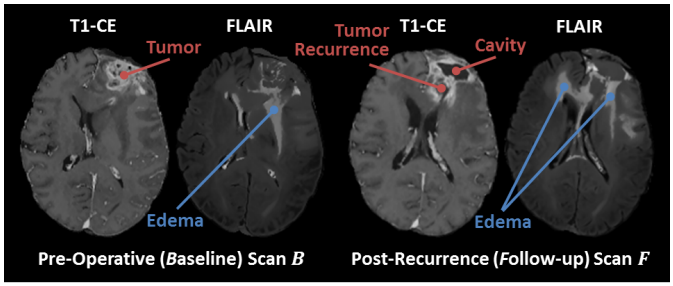


Fig. 1. An example of the pre-operative scan B and the corresponding post-recurrence scan F. The pre-operative scan B clearly shows the tumor in the T1-CE scan and edema in FLAIR scan. Edema is also clearly visible in the FLAIR scan of the post-recurrence scan F. The T1-CE of F now shows resection cavity and tumor recurrence.

much more difficult problem of registering an atlas of a healthy population to scans showing pathology. Alternatively, one can directly register the scans using state-of-the-art intensity- or feature-based registration methods [18]–[20]. However, these approaches do not explicitly account for pathologies and therefore may produce unreasonable correspondences in these areas. *PORTR* instead makes use of the fact that both scans are from the same patient and explicitly models constraints enforced by the pathology in each scan.

PORTR determines the optimal deformation between two scans by finding the minimum of an energy function, which is based on the concept of symmetric registration [19], [21]–[24]. This energy function is not only comprised of image-based correspondences and smoothness constraints as customary for other registration methods, but also includes pathological information. The pathological information is inferred from the results of two segmenters that are targeted to each scan. Specifically, we develop a new method for segmenting post-recurrence scans, which generally consist of resection cavities after brain surgery and multiple tumor recurrences. For the pre-operative scans, we adapt the segmenter by Gooya *et al.* [16] to outline a single brain glioma which causes a large mass effect on healthy tissue. The resulting segmentations of both scans are a central component in the definition of the image and the shape-based correspondence terms within our symmetric registration framework. Determining the minimum within this framework is difficult as the function contains many local minima. We deal with these difficulties by combining discrete and continuous optimizations. The discrete optimization method finds the optimal solution in a coarse solution space. The continuous optimization method locally improves this solution in a finer solution space. We measure the accuracy of *PORTR* on 24 subjects. The results indicate that the proposed method outperforms Avants *et al.* [19] and Ou *et al.* [20], two examples of the state-of-the-art in general-purpose registration methods.

II. A Deformable Registration Framework for Tumor Scans

We now describe *PORTR*, which aligns the pre-operative scan (baseline, denoted as B) with the post-recurrence scan (follow-up, denoted as F) of the same subject. As outlined in Fig. 2, our approach first applies atlas-based segmenters to

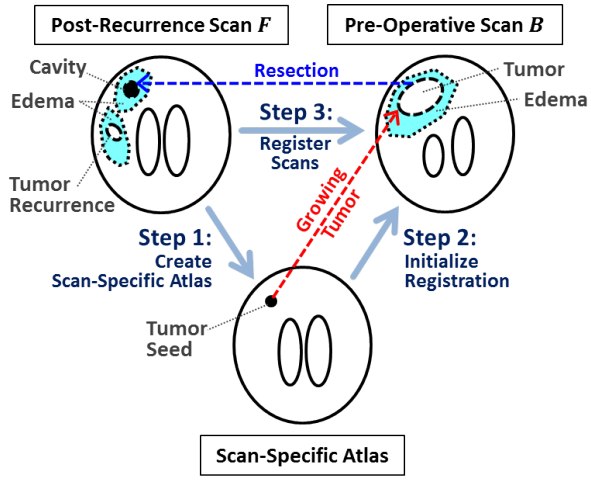


Fig. 2. Our deformable registration framework consists of three steps. We first create a scan-specific atlas for F by segmenting the scan (Step 1). We then jointly register this atlas to B and segment the scan, which provides an initial mapping between F and B (Step 2). Finally, we register the scans using the results from the previous two steps (Step 3).

F (Step 1) and B (Step 2) to extract pathological information needed to register the scans (Step 3). Our analysis starts with F instead of B as the glioma shown in B is surgically removed in F. Thus, the healthy tissue of F is not impacted by the large mass effect of that glioma [25], [26] so that segmenters guided by atlases of healthy populations are generally easier adapted to F than B. We do so in Step 1 by exploring a new multi-tumor model for modifying the atlas of the healthy population to scan F. We then interpret the results of Step 1 as a scan-specific atlas for Step 2 guiding the joint atlas registration and segmentation of the pre-operative scan B [16]. The registration of Step 2 initializes the registration between F and B of Step 3. For Step 3, we propose a new probabilistic registration framework coupling the results of the previous steps with the image-based correspondences between F and B.

A. Step 1: Segmentation of Post-Recurrence Scan

The goal of this step is to compute probabilities regarding the presence of healthy tissue and pathology within the post-recurrence scan F. We do so by simultaneously registering an atlas to scan F and segmenting the anatomy in that scan. The atlas provides spatial information about the structures of interest, which is needed to distinguish between the similar MR intensity patterns of healthy tissue and pathology. In the remainder of this section, we first provide a simple model for transforming an atlas of a healthy population to one including pathological information specific to scan F. We then integrate the atlas into our Bayesian approach for joint segmentation and registration of the post-recurrence scan.

Before describing our probabilistic model in further detail, we introduce the following nomenclature: $\mathcal{T}$ denotes the label map across the image domain Ω_F of scan F. The possible label t of $\mathcal{T}$ at a specific location $\mathbf{x} \in \Omega_F$ is WM (white matter), GM (gray matter), CSF (cerebrospinal fluid), ED (edema), and TU (tumor), which includes enhanced tumor,

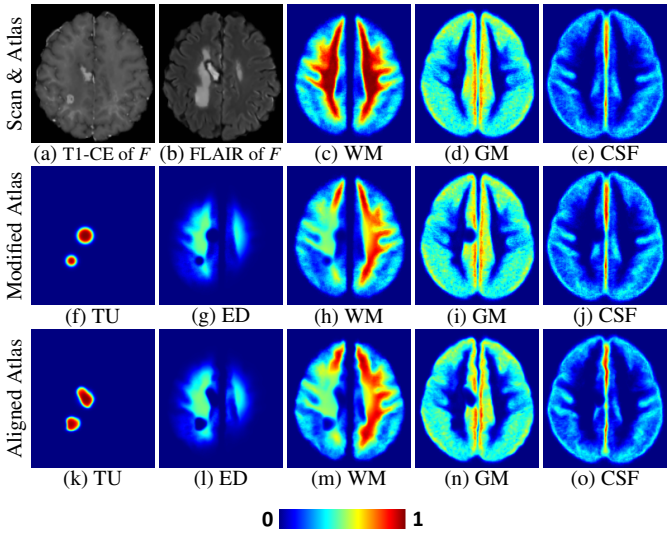


Fig. 3. The atlas created by Step 1. The figure shows the post-recurrence scan F in (a) and (b), and a probabilistic atlas of a healthy population in (c)-(e). (f)-(j) are the spatial probabilities $p_F(\mathcal{T}_t)$ based on our multi-tumor model applied to the healthy atlas. (k)-(o) shows the spatial probabilities aligned to the scan. Those spatial distributions better fit to the scan than the original ones (f)-(j).

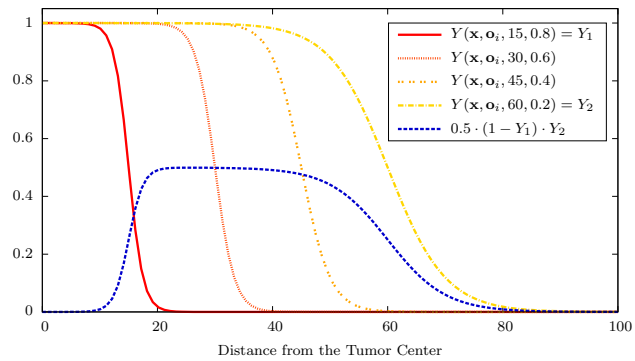


Fig. 4. Radial profiles of spatial probability functions Y with varying parameters. As Y is radially symmetric, we set the x-axis as the distance from the tumor center ($\mathbf{o}_i$ is located at zero). The four sample curves for Y are based on different parameter settings for $r_i \in \{15, 30, 45, 60\}$ and $a \in \{0.8, 0.6, 0.4, 0.2\}$. The plot in blue is an edema model combining graphs labeled as Y_1 and Y_2. The graphs show the curves start to slope further away from the origin the larger r_i is and their slope steepen the larger a is.

necrosis and cavity. For short, we denote this as '$\mathcal{T}_t|\mathbf{x}$' instead of '$\mathcal{T} = t|\mathbf{x}$'. p_A is the probabilistic brain tissue atlas of a healthy population, which we assume to be affinely aligned with F. Let $\Theta_A \triangleq \{WM, GM, CSF\}$ be the labels of the healthy tissue types then $p_A(\mathcal{T}_t|\mathbf{x})$ is the probability of tissue $t \in \Theta_A$ being present at location $\mathbf{x} \in \Omega_F$ (see Fig. 3 (c)-(e)). Similarly, p_F corresponds to the probabilistic model associated with scan F, $\Theta \triangleq \{WM, GM, CSF, ED, TU\}$ are the labels of all possible tissue types on p_F, and the spatial probability $p_F(\mathcal{T}_t|\mathbf{x})$ is the conditional probability of tissue $t \in \Theta$ being present at location $\mathbf{x} \in \Omega_F$.

We now define $p_F(\mathcal{T}_t|\mathbf{x})$ for each $t \in \Theta$ by combining p_A with a simple model for pathology. This model is based on the empirical observation that the post-recurrence scan F generally shows multiple small tumor recurrences, resection cavities, and edema. These pathologies do not cause a large mass effect on the healthy tissue in general. Our spatial probabilities $p_F(\mathcal{T}_t|\mathbf{x})$ are therefore based on the simplifying assumptions that the mass effect of the pathologies on healthy tissue can be ignored and that each pathological region is contained within a relatively small sphere. If we assume F shows M tumors then each tumor $i \in \{1, \dots M\}$, which we loosely use for tumor recurrences and resection cavities, can thus be characterized by its center location $\mathbf{o}_i \in \Omega_F$ and its size or radius r_i. We manually set $\mathbf{o}_i$ and r_i so that the resulting sphere encompasses the abnormal region as shown on T1-CE scans. Based on the previous assumptions, we model its corresponding spatial probability via the generalized logistic function [27]:

$$Y(\mathbf{x}, \mathbf{o}_i, r_i, a) \triangleq \frac{1}{1 + \exp\{a \cdot (||\mathbf{x} - \mathbf{o}_i|| - r_i)\}} , \quad (1)$$

where $||\cdot||$ is the ℓ^2-norm and a controls the steepness of the function. Fig. 4 plots the radial profile for Eq. (1) with different parameters. The graph shows that the curves start to slope further away from the origin as r_i gets larger and their slope steepens as a gets larger. This slope represents the diffusion of the tumor into healthy tissue. For simplicity, we set $a = a_{TU}$ uniformly across all tumors. The spatial probability $p_F(\mathcal{T}_{TU}|\mathbf{x})$ across all tumors is then defined by the maximum value of the M individual spatial probabilities Y as follows:

$$p_F(\mathcal{T}_{TU}|\mathbf{x}) \triangleq \max_i \Big\{ Y(\mathbf{x}, \mathbf{o}_i, r_i, a_{TU}) \Big\} . \quad (2)$$

We use the maximum across the spatial distributions of all tumors as the function preserves the probabilistic profile of individual tumors (assuming they are well spaced) and ensures the range of $p_F(\mathcal{T}_{TU}|\mathbf{x})$ to be within $[0, 1]$.

The spatial probability of edema is based on the assumption that edema is in close proximity of tumors and its signal strength decays smoothly as the distance from the tumor center increases. Furthermore, edema is contained inside white matter and should be defined in relation to tumor. In other words, the more probable the presence of tumor, the less probable edema should be. These assumptions are summarized in the following definition of the spatial probability of edema:

$$p_F(\mathcal{T}_{ED}|\mathbf{x}) \triangleq 0.5 \cdot p_A(\mathcal{T}_{WM}|\mathbf{x}) \cdot \big(1 - p_F(\mathcal{T}_{TU}|\mathbf{x})\big) \cdot \max_i \Big\{ Y(\mathbf{x}, \mathbf{o}_i, b_{ED} \cdot r_i, a_{ED}) \Big\} . \quad (3)$$

We set $a_{ED} < a_{TU}$, which means edema is more dispersed than tumor. b_{ED} defines the area of edema with respect to the ith pathological region. Also, the factor 0.5 ensures that our atlas does not favor edema over WM (or vice versa) in areas of $Y \approx 1$ in Eq. (3). Fig. 4 shows the radial profile for $p_F(\mathcal{T}_{TU}|\mathbf{x})$ (red line) and $p_F(\mathcal{T}_{ED}|\mathbf{x})$ (blue line) assuming a single tumor is present and its size is $r_i = 15$. One can see that the probability of edema is close to zero inside the tumor, increases to 0.5, and then it smoothly decays to zero again.

Next, we model the spatial probabilities of the healthy tissue classes. We combine the atlas p_A with $p_F(\mathcal{T}_{TU}|\mathbf{x})$ based on the observation that in areas where $p_F(\mathcal{T}_{TU}|\mathbf{x})$ is relatively

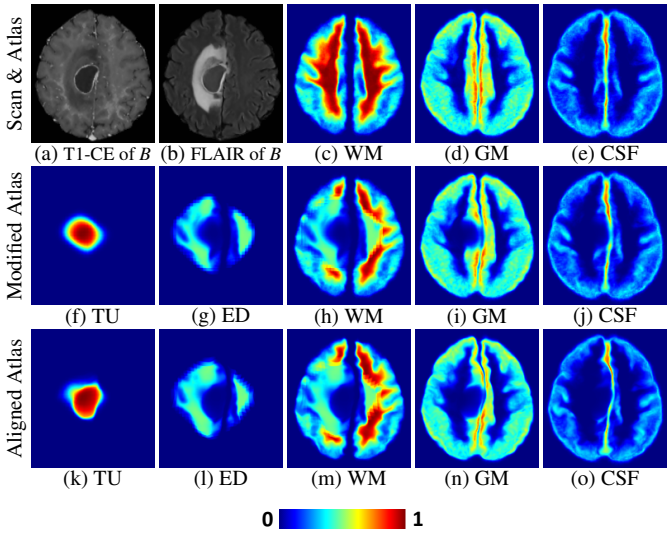


Fig. 5. An example of spatial probabilities in Step 2. We show the pre-operative scan B in (a) and (b) and scan-specific atlas in (c)-(e). In (f)-(j), we show spatial probabilities $p_B(\mathcal{T}_t)$ obtained by applying the tumor growth model on scan-specific atlas. (k)-(o) shows the spatial probabilities aligned to the scan. This atlas now fits well to healthy tissue and pathological regions shown in (a) and (b).

large, the probability of healthy tissue should be small. For GM and CSF, this observation is reflected by the following product ($t \in \{GM, CSF\}$):

$$p_F(\mathcal{T}_t|\mathbf{x}) \triangleq p_A(\mathcal{T}_t|\mathbf{x}) \cdot \big(1 - p_F(\mathcal{T}_{TU}|\mathbf{x})\big) \ .$$

For WM, we take the complement of spatial probabilities of the other labels:

$$p_F(\mathcal{T}_{WM}|\mathbf{x}) \triangleq 1 - \big\{p_F(\mathcal{T}_{TU}|\mathbf{x}) + p_F(\mathcal{T}_{ED}|\mathbf{x}) + p_F(\mathcal{T}_{GM}|\mathbf{x}) + p_F(\mathcal{T}_{CSF}|\mathbf{x})\big\} \ .$$

Note that $p_F(\mathcal{T}_{WM}|\mathbf{x})$ is always non-negative. Fig. 3 (f)-(j) shows an example of our spatial probabilities. The spatial probability for edema is high in close proximity of tumors. Furthermore, spatial probabilities for WM, GM and CSF are decreased in comparison to their values in the healthy atlas in areas where tumor or edema appears to be present.

Having defined the spatial probabilities $p_F(\mathcal{T}_t|\mathbf{x})$, we now describe our approach for computing the posterior probabilities of all tissue types. For $p_F(\mathcal{T}_t|\mathbf{x})$ to be informative it needs to match F. Fig. 3 (f)-(j) shows that this is generally not the case. For example, $p_F(\mathcal{T}_t|\mathbf{x})$ implies a high probability of tumor outside the pathology shown in F. We address this issue by jointly computing posterior probabilities and registering $p_F(\mathcal{T}_t|\mathbf{x})$ to F.

One of the parameters of our joint registration and segmentation model is $\boldsymbol{h}_F$, the 'voxel wise mapping' from the Ω_F to the atlas space. A voxel-wise mapping projects a voxel from the source to target space according to the 3D vector stored in the underlying deformation map at that location. The second set of parameters are the tissue specific mean and covariances $\boldsymbol{\Phi}_F$, which define the multivariate Gaussian of the image likelihood $p_F(F|\mathcal{T}_t, \boldsymbol{\Phi}_F, \mathbf{x})$. The joint registration and segmentation problem is then defined via the following optimization problem

$$\{\boldsymbol{h}_F^*, \boldsymbol{\Phi}_F^*\} \triangleq \arg\max_{\boldsymbol{h}_F, \boldsymbol{\Phi}_F} \Big(\prod_{\mathbf{x} \in \Omega_F} \sum_{t \in \Theta} p_F(\mathcal{T}_t|\boldsymbol{h}_F(\mathbf{x})) \cdot p_F(F|\mathcal{T}_t, \boldsymbol{\Phi}_F, \mathbf{x}) \Big) \ . \quad (4)$$

We obtain $\boldsymbol{h}_F^*$ and $\boldsymbol{\Phi}_F^*$ via an implementation of the Expectation-Maximization (EM) algorithm [28]. The details of this implementation are provided in Appendix A. Then we define the posteriors for the post-recurrences scan with respect to the anatomy $t \in \Theta$ as

$$\begin{aligned} p_F(\mathcal{T}_t|F, \mathbf{x}) &\triangleq p_F(\mathcal{T}_t|F, \boldsymbol{h}_F^*, \boldsymbol{\Phi}_F^*, \mathbf{x}) \\ &\propto p_F\big(\mathcal{T}_t|\boldsymbol{h}_F^*(\mathbf{x})\big) \cdot p_F(F|\mathcal{T}_t, \boldsymbol{\Phi}_F^*, \mathbf{x}) \ . \end{aligned} \quad (5)$$

Fig. 3 (k)-(o) show the aligned spatial probability $p_F(\mathcal{T}_t|\boldsymbol{h}_F^*(\mathbf{x}))$ obtained by applying $\boldsymbol{h}_F^*$ to the original spatial probability $p_F(\mathcal{T}_t|\mathbf{x})$ of Fig. 3 (f)-(j). The probability for each tissue is transformed to match the scan. The aligned spatial probability of tumor (k) correlates now very well with the scan compared to its original (f).

B. Step 2: Segmentation of Pre-Operative Scan

The goal of this step is to segment pathological regions from the pre-operative scan B and to provide a rough estimation of the deformation between F and B. We achieve this goal via the joint segmentation and registration approach by Gooya *et al.* [16], called *GLISTR*. *GLISTR* explicitly models the generally large mass effects on healthy tissue caused by brain glioma in scan B. We now describe the integration of the results of Step 1 into this approach and provide a brief review of the method.

Similar to Step 1, *GLISTR* uses the EM algorithm to jointly register an atlas to the scan B and segment the scan B into healthy tissue and pathological regions. For our specific application, we replace the atlas of a healthy population proposed in [16] with the scan-specific atlas defined for tissue $t \in \Theta_A$ as

$$p_S(\mathcal{T}_t|\mathbf{x}) \triangleq p_A(\mathcal{T}_t|\boldsymbol{h}_F^*(\mathbf{x})) \ , \quad (6)$$

where the mapping $\boldsymbol{h}_F^*$ was defined according to Eq. (4) in Step 1. Fig. 5 (c)-(e) shows an example of the scan-specific atlas. This atlas is not affected by pathology and is aligned with F (see Fig. 3 (a) and (b)).

By registering this atlas to scan B, the method also approximates the mapping between F and B. We note that we could have also based the atlas on the posteriors of Eq. (5) instead of $p_A(\mathcal{T}_t|\boldsymbol{h}_F^*(\mathbf{x}))$. Compared to $p_A(\mathcal{T}_t|\boldsymbol{h}_F^*(\mathbf{x}))$, the posteriors have generally higher certainty about the presence or absence of healthy tissue throughout the image domain. The higher certainty causes the method to have less flexibility in registering p_S to B. In practice, this makes the registration problem more difficult causing the method to be less stable than when using our proposed atlas.

The remainder of this section provides a brief overview of how *GLISTR* simultaneously models tumor growth in the atlas space, registers the corresponding atlas to scan B, and segments B. We denote with p_B the probabilistic model specific to the scan B and Ω_B as the space of the pre-operative

scan. As in the previous step, $\mathcal{T}$ represents the label map that at each image location is assigned to the labels of Θ. The approach adapts p_S to B by simulating tumor growth on Ω_F via the diffusion-reaction-advection model by Hogea *et al.* [29]. Given the parameters $\mathbf{q}$, which contain the seed location of the tumor, the model produces the tumor probability $p_B(\mathcal{T}_{TU}|\mathbf{q},\mathbf{x})$ and the voxel-wise mapping $\mathbf{u}$ according to this tumor probability. We manually set the seed in the center of the tumor. The approach then combines $p_B(\mathcal{T}_{TU}|\mathbf{q},\mathbf{x})$ and $\mathbf{u}$ with the atlas p_S. Now, the spatial probabilities for GM and CSF ($t \in \{GM, CSF\}$) are defined as

$$p_B(\mathcal{T}_t|\mathbf{q},\mathbf{x}) \triangleq p_S(\mathcal{T}_t|\mathbf{u}(\mathbf{x})) \cdot \big(1 - p_B(\mathcal{T}_{TU}|\mathbf{q},\mathbf{x})\big) \ . \quad (7)$$

Unlike the spatial probability of edema in Step 1 (Eq. (3)), *GLISTR* models the close proximity of edema to tumor via the Heaviside function $H(\cdot)$, ($H(a) = 0$ for $a \leq 0$ and $H(a) = 1$ for $a > 0$) resulting in

$$p_B(\mathcal{T}_{ED}|\mathbf{q},\mathbf{x}) \triangleq 0.5 \cdot p_S(\mathcal{T}_{WM}|\mathbf{u}(\mathbf{x})) \cdot \big(1 - p_B(\mathcal{T}_{TU}|\mathbf{q},\mathbf{x})\big) \cdot H\big(p_B(\mathcal{T}_{TU}|\mathbf{q},\mathbf{x})\big) \ , \quad (8)$$

where we multiply 0.5 in order to avoid preference of edema over WM (or vice versa) as in Eq. (3). The Heaviside function explicitly confines edema to the region inferred from the outcome of the tumor growth model represented by $p_B(\mathcal{T}_{TU}|\mathbf{q},\mathbf{x})$. This region generally encompasses edema as the tumor growth model accounts for diffusion into healthy tissue. Note that Step 1 did not include a dynamic tumor model. We instead modeled diffusion through the logistic function with a fixed slope (Eq. (1)), which we then used to define the edema region. The spatial probability for WM is defined by the complement of spatial probabilities of the other labels:

$$p_B(\mathcal{T}_{WM}|\mathbf{q},\mathbf{x}) \triangleq 1 - \big\{p_B(\mathcal{T}_{TU}|\mathbf{q},\mathbf{x}) + p_B(\mathcal{T}_{ED}|\mathbf{q},\mathbf{x}) + p_B(\mathcal{T}_{GM}|\mathbf{q},\mathbf{x}) + p_B(\mathcal{T}_{CSF}|\mathbf{q},\mathbf{x})\big\} \ .$$

The EM algorithm determines the optimal parameters of this model, which are the tumor parameters $\mathbf{q}^*$, the voxel-wise mapping $\boldsymbol{h}_B^*$ of the posteriors of B to the scan-specific atlas, and the Gaussian intensity distribution parameters $\boldsymbol{\Phi}_B^*$. The posterior of structure $t \in \Theta$ for the pre-operative scan B is then defined as

$$\begin{aligned} p_B(\mathcal{T}_t|B,\mathbf{x}) &\triangleq p_B(\mathcal{T}_t|B,\mathbf{q}^*,\boldsymbol{h}_B^*,\boldsymbol{\Phi}_B^*,\mathbf{x}) \\ &\propto p_B(\mathcal{T}_t|\mathbf{q}^*,\boldsymbol{h}_B^*(\mathbf{x})) \cdot p_B(B|\mathcal{T}_t,\boldsymbol{\Phi}_B^*,\mathbf{x}) \ . \end{aligned} \quad (9)$$

Let $\mathbf{u}^*$ represent the mapping according to the tumor growth model, which is parameterized by $\mathbf{q}^*$. We approximate the mapping between scan B and F by concatenating '$\circ$' two mappings $\boldsymbol{h}_B^*$ and $\mathbf{u}^*$:

$$\boldsymbol{f}_{BF}^0 \triangleq \mathbf{u}^* \circ \boldsymbol{h}_B^* \ . \quad (10)$$

Note that the voxel-wise mapping $\boldsymbol{f}_{BF}^0$ approximates the alignment from B to F as we use p_S instead of image and pathological information for F.

Fig. 5 (f)-(j) shows an example of spatial probabilities $p_B(\mathcal{T}_t|\mathbf{x})$. The spatial probability of tumor (f) covers the tumor shown in (a) and (b), the one of edema (g) is neighboring the tumor, and probabilities of healthy tissues are displaced by mass effect of the tumor. Fig. 5 (k)-(o) shows aligned spatial probabilities $p_B(\mathcal{T}_t|\boldsymbol{h}_B^*(\mathbf{x}))$ which better seem to fit the scan than those of Fig. 5 (f)-(j).

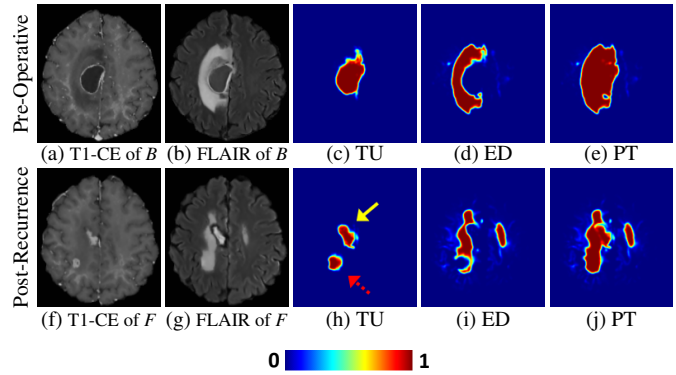


Fig. 6. An example of posteriors estimated by Step 1 and 2. For the pre-operative scan B shown in (a) and (b), the posteriors of (c) tumor and (d) edema are obtained by Step 2. For the post-recurrence scan F shown in (f) and (g), posteriors of (h) tumor and (i) edema are obtained by Step 1. In (h), the yellow arrow marks the regions for cavity and the red arrow marks the region for tumor recurrence. The probabilities for pathological regions (e) $p_{B,PT}$ and (j) $p_{F,PT}$ are defined as the sum of the posteriors of tumor and edema. In (e) and (j), the regions indicating high probability of tumor being present are matched well with pathological regions in B and F, respectively. Thus, those two maps are well suited for masking out pathological regions in the matching cost function.

C. Step 3: Deformable Registration of Pre-Operative and Post-Recurrence Scans

The goal is now to register the pre-operative scan B and the post-recurrence scan F to accurately match the intensities of the non-pathological regions while simultaneously inferring reasonable deformations for the pathology based on the posterior probabilities of the previous two steps. We do so by applying the concept of symmetric registration [19], [21]–[24] to our scenario. The idea behind the symmetric registration is not to favor either scan by matching both scans to a 'center coordinate system'. The mapping between the two scans is now essentially split into two. This splitting allows us to reliably determine the large deformations between the scans. To determine the optimal mapping, we apply a hybrid optimization method combining discrete and continuous optimizations. It allows us to refine the global solution of the coarse search space determined by the discrete optimization with the local search by the continuous optimization.

1) Symmetric Registration Framework: Let Ω_C be the center coordinate system, $\boldsymbol{f}_{CB} : \Omega_C \to \Omega_B$ the diffeomorphic mapping from Ω_C to Ω_B, $\boldsymbol{f}_{CF} : \Omega_C \to \Omega_F$ the diffeomorphic mapping from Ω_C to Ω_F, and '$\circ$' concatenates two mappings. Then

$$\boldsymbol{f}_{BF} \triangleq \boldsymbol{f}_{CF} \circ (\boldsymbol{f}_{CB})^{-1} \quad (11)$$

maps B to F. The solution to our symmetric registration problem $\{\boldsymbol{f}_{CB}^*, \boldsymbol{f}_{CF}^*\}$ minimizes an energy function $E(\cdot)$

$$\{\boldsymbol{f}_{CB}^*, \boldsymbol{f}_{CF}^*\} = \arg \min_{\boldsymbol{f}_{CB},\boldsymbol{f}_{CF}} E\big(\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}; B, F, p_B, p_F\big) \ . \quad (12)$$

$E(\cdot)$ encodes the relationship between Ω_C, Ω_B and Ω_F by a *correspondence term* E_C measuring the agreement between image patches of B and F, a *pathology term* E_P capturing the overlap between posteriors of tumors in both scans, and a *smoothness term* E_S enforcing consistency across the deformation fields. Thus the energy function is of the form

$$\begin{aligned} E(\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}; B, F, p_B, p_F) \triangleq \\ \lambda_D \cdot E_C(\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}; B, F, p_B, p_F) \\ + \lambda_D \cdot \lambda_P \cdot E_P(\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}; p_B, p_F) \\ + E_S(\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}) \ , \end{aligned} \tag{13}$$

where λ_D is a weight of data terms (E_C and E_P) compared to E_S and λ_P is a weight of E_P among data terms. The remainder of this section describes the three terms in further detail.

The *correspondence term* E_C is based on Normalized Cross-Correlation (NCC) [30] to measures the image similarity of healthy tissue between the aligned baseline scan $B \circ \boldsymbol{f}_{CB}$ and follow-up scan $F \circ \boldsymbol{f}_{CF}$ in Ω_C. We choose NCC as it is often used for intra-subject registrations [18]. In our case, each scan consists of N co-registered, multi-modal images (e.g. T1, T1-CE, T2, and FLAIR) and B_i (or F_i) denotes the ith image of B (or F). The NCC of these multi-modal images at voxel $\mathbf{x} \in \Omega_C$ is the mean NCC score across the modalities

$$D_{NCC}(B \circ \boldsymbol{f}_{CB}, F \circ \boldsymbol{f}_{CF}, \mathbf{x}) \triangleq \frac{1}{N} \sum_{i=1}^{N} \frac{\langle \overline{B_i}, \overline{F_i} \rangle}{\sqrt{\langle \overline{B_i}, \overline{B_i} \rangle \cdot \langle \overline{F_i}, \overline{F_i} \rangle}} \ , \tag{14}$$

with $\langle \cdot, \cdot \rangle$ being the inner product of aligned and intensity-corrected patches $\overline{B_i}$ and $\overline{F_i}$. To compute $\overline{B_i}$ and $\overline{F_i}$, we define the region of the patch $\mathcal{R}(\mathbf{x}) \subset \Omega_C$ centered around $\mathbf{x}$ and compute the mean intensity value $m(\mathbf{x})$ for that patch. Then

$$\overline{I_i} \triangleq \left\{ I_i(\boldsymbol{f}_{CI}(\mathbf{y})) - m(\mathbf{x}) \mid \mathbf{y} \in \mathcal{R}(\mathbf{x}) \right\} ,$$

for each $I \in \{B, F\}$. To confine the correspondence term $E_C(\cdot)$ to healthy tissue, we incorporate the probability for pathological regions in this term. As the pathology is indicated by tumor and edema, the probability for pathological regions $p_{I,PT}$ in each scan $I \in \{B, F\}$ is defined as the sum of the posteriors of tumor and edema, i.e.

$$p_{I,PT}(\mathbf{x}) \triangleq p_I(\mathcal{T}_{TU} | I, \mathbf{x}) + p_I(\mathcal{T}_{ED} | I, \mathbf{x}) \ .$$

As shown in Fig. 6 (e) and (j), $p_{B,PT}$ and $p_{F,PT}$ are good indicators for pathology. Now we know that pathology creates image patterns that are scan specific and thus unreliable for image matching. The opposite is true for healthy tissue. One way to reflect this observation in E_C is to use these probabilities as saliency information weighing D_{NCC} more in healthy regions and less in pathological regions. The following definition of the correspondence terms does exactly that

$$\begin{aligned} E_C(\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}; B, F, p_B, p_F) \triangleq \\ \int_{\mathbf{x} \in \Omega_C} \left\{ 1 - p_{B,PT}(\boldsymbol{f}_{CB}(\mathbf{x})) \right\} \cdot \left\{ 1 - p_{F,PT}(\boldsymbol{f}_{CF}(\mathbf{x})) \right\} \\ \cdot D_{NCC}(B \circ \boldsymbol{f}_{CB}, F \circ \boldsymbol{f}_{CF}, \mathbf{x}) \ d\mathbf{x} \ . \end{aligned} \tag{15}$$

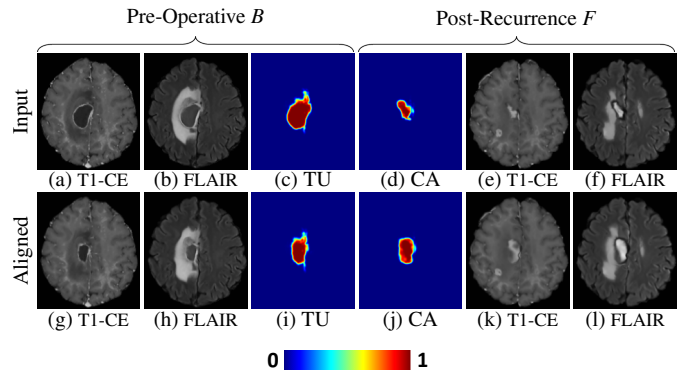


Fig. 7. An example of scans and posteriors aligned by Step 3. The upper row shows input scans and their estimated posteriors of (c) tumor and (d) cavity. The posterior of cavity (d) is obtained from the posterior of tumor shown in Fig. 6 (h) using Eq. (16). The lower row shows the aligned scans and posteriors. Specifically, (g)-(i) are warped from (a)-(c) using $\boldsymbol{f}^*_{CB}$ and (j)-(l) are warped from (d)-(f) using $\boldsymbol{f}^*_{CF}$. Now the tumor nicely matches the cavity region.

The *pathology term* E_P measures the overlap between the tumor in B and the resection cavity in F. Similar to E_C, we use the posterior of tumor of Step 2 to indicate the tumor region in B. To flag the cavity region in F, we separate the seeds used for the posterior of tumor of Step 1 in those that are associated with cavities versus those with tumor recurrences. We then set the indicator variable $\mathbf{1}_{CA}(\mathbf{x})$ at voxel $\mathbf{x} \in \Omega_F$ to one if the index $i(\mathbf{x}) = \arg\max_i \{Y(\mathbf{x}, \mathbf{o}_i, r_i, a)\}$ used in Eq. (2) corresponds to seeds for the cavity, and otherwise set to zero. Then the posterior probability of cavity is

$$p_{F,CA}(\mathbf{x}) \triangleq p_F(\mathcal{T}_{TU} | F, \mathbf{x}) \cdot \mathbf{1}_{CA}(\boldsymbol{h}^*_F(\mathbf{x})) \ . \tag{16}$$

Fig. 7 (d) shows an example of $p_{F,CA}$ obtained from Fig. 6 (h) using Eq. (16). Only the cavity region of Fig. 6 (h) is correctly selected in Fig. 7 (d).

We now define the pathology term E_P in such a way that it penalizes mismatches between the aligned posteriors of the tumor in B and of the cavity in F, i.e. we measure the squared ℓ^2-norm between the aligned $p_{B,TU}$ and $p_{F,CA}$ [31], [32]

$$\begin{aligned} E_P(\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}; p_B, p_F) \triangleq \\ \int_{\mathbf{x} \in \Omega_C} \left\{ p_{B,TU}(\boldsymbol{f}_{CB}(\mathbf{x})) - p_{F,CA}(\boldsymbol{f}_{CF}(\mathbf{x})) \right\}^2 d\mathbf{x} \ . \end{aligned} \tag{17}$$

In our experiments, this term performed slightly better than information theoretic measures such as the Jensen-Shannon (JS) divergence [33]. This is not surprising for shape alignment according to Wang *et al.* [34]. More importantly, the above term is more efficient to compute than JS. During optimizations, E_P leads tumor and cavity regions to correspond to each other which is difficult to do based on image-based correspondences. Fig. 7 (c) and (d) show the tumor and cavity posteriors, and Fig. 7 (i) and (j) show the corresponding regions with (c) and (d) transformed to have similar shapes after optimizations.

The third and final term of $E(\cdot)$ is the *smoothness term* E_S, which penalizes discontinuities in both $\boldsymbol{f}_{CB}$ and $\boldsymbol{f}_{CF}$. E_S measures the smoothness of the mappings $\boldsymbol{f}_{CB}$ and $\boldsymbol{f}_{CF}$ via the Tikhonov operator L [35]–[37]. Let c_i be a non-negative

Input: $\boldsymbol{f}^0_{BF}, B, F, p_B, p_F$
Output: $\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}$
// (Initialize solution)
set $\boldsymbol{f}_{CB} \leftarrow \left(\mathrm{Id} + \frac{\boldsymbol{f}^0_{BF} - \mathrm{Id}}{2}\right)^{-1}$ and $\boldsymbol{f}_{CF} \leftarrow \left(\mathrm{Id} + \frac{\boldsymbol{f}^{0\ -1}_{BF} - \mathrm{Id}}{2}\right)^{-1}$
for *coarse* **to** *fine scale* **do**
  resample $\{B, F, p_B, p_F, \boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}\}$ to current scale
  // (Apply discrete optimization)
  set $B' \leftarrow B \circ \boldsymbol{f}_{CB}$, $p_{B'} \leftarrow p_B \circ \boldsymbol{f}_{CB}$, $\boldsymbol{f}_{B'B} \leftarrow \boldsymbol{f}_{CB}$
    $F' \leftarrow F \circ \boldsymbol{f}_{CF}$, $p_{F'} \leftarrow p_F \circ \boldsymbol{f}_{CF}$, $\boldsymbol{f}_{F'F} \leftarrow \boldsymbol{f}_{CF}$
  calculate $\{\boldsymbol{f}_{CB'}, \boldsymbol{f}_{CF'}\}$ using Eq. (20) and Eq. (21)
  set $\boldsymbol{f}_{CB} \leftarrow \boldsymbol{f}_{B'B} \circ \boldsymbol{f}_{CB'}$ and $\boldsymbol{f}_{CF} \leftarrow \boldsymbol{f}_{F'F} \circ \boldsymbol{f}_{CF'}$
  // (Apply continuous optimization)
  repeat
    set $B' \leftarrow B \circ \boldsymbol{f}_{CB}$, $p_{B'} \leftarrow p_B \circ \boldsymbol{f}_{CB}$, $\boldsymbol{f}_{B'B} \leftarrow \boldsymbol{f}_{CB}$
      $F' \leftarrow F \circ \boldsymbol{f}_{CF}$, $p_{F'} \leftarrow p_F \circ \boldsymbol{f}_{CF}$, $\boldsymbol{f}_{F'F} \leftarrow \boldsymbol{f}_{CF}$
    calculate $\{\boldsymbol{f}_{CB'}, \boldsymbol{f}_{CF'}\}$ using Eq. (23)
    set $\boldsymbol{f}_{CB} \leftarrow \boldsymbol{f}_{B'B} \circ \boldsymbol{f}_{CB'}$ and $\boldsymbol{f}_{CF} \leftarrow \boldsymbol{f}_{F'F} \circ \boldsymbol{f}_{CF'}$
  until *convergence*
end

Algorithm 1: Our hybrid optimization method. We initialize the solution using $\boldsymbol{f}^0_{BF}$ and then update the mappings $\{\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}\}$ from coarse to fine scales sequentially executing discrete and continuous optimizations for each image resolution.

constant and $\mathrm{Id}(\mathbf{x}) = \mathbf{x}$ be the identity mapping, then the Tikhonov operator of a mapping $\boldsymbol{f}$ is

$$L\boldsymbol{f} \triangleq \sum_{i=0}^{\infty} c_i \cdot \left|\nabla^i(\boldsymbol{f} - \mathrm{Id})\right|^2 , \tag{18}$$

where $\nabla^i(\boldsymbol{f} - \mathrm{Id})$ is the ith order derivative of the displacement field $(\boldsymbol{f} - \mathrm{Id})$. The smoothness term is now defined as

$$E_S(\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}) \triangleq \int_{\mathbf{x}\in\Omega_C} \Big\{ L\boldsymbol{f}_{CB}(\mathbf{x}) + L\boldsymbol{f}_{CF}(\mathbf{x}) \Big\}\, d\mathbf{x} . \tag{19}$$

This regularizer favors smooth deformations as it penalizes magnitude of higher order derivatives [36]. We note that our method is not specific to the Tikhonov operator so that any other operator penalizing discontinuities in the displacement field could be used at this point.

This completes our definition of the energy function E, whose minimum (see Eq. (12)) defines our solution for determining the mappings $\boldsymbol{f}_{CB}$ and $\boldsymbol{f}_{CF}$.

2) Hybrid Optimization Method: We determine the minimum of Eq. (12) via a hybrid approach combining discrete with continuous optimization, a concept recently explored in optical flow estimations [38], [39]. Discrete optimizations generally determine the global minimum (or a strong local minimum) but do so with respect to a limited search space. On the other hand, continuous optimization methods search in a much richer solution space but often get trapped in local minima causing them to be sensitive towards their initialization. To take advantage of both approaches, we first apply our discrete optimization method and use those results as initialization of the continuous optimization method.

As outlined in Algorithm 1, we initialize our algorithm by 'splitting' the initial displacement $\boldsymbol{f}^0_{BF}$ of Eq. (10) in half:

$$\boldsymbol{f}_{CB} \leftarrow \left(\mathrm{Id}+\frac{\boldsymbol{f}^0_{BF} - \mathrm{Id}}{2}\right)^{-1},\ \boldsymbol{f}_{CF} \leftarrow \left(\mathrm{Id}+\frac{\boldsymbol{f}^{0\ -1}_{BF} - \mathrm{Id}}{2}\right)^{-1}.$$

We note that this initialization is one of many schemes that fulfill Eq. (11), i.e. the requirement of our symmetric registration framework. We choose this specific one due to its simplicity.

We then successively apply our discrete and continuous optimization methods based on the coarse-to-fine scheme [40]. At each iteration, our algorithm determines the deformations maps $\{\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}\}$ that minimize the energy function of Eq. (13) for the resolution associated with this iteration. In the ideal case, the resulting intermediate images $B' \triangleq B \circ \boldsymbol{f}_{CB}$ and $F' \triangleq F \circ \boldsymbol{f}_{CF}$ are equal on the healthy tissue with respect to this resolution. The remainder of this section describes our discrete and continuous optimizations in further detail.

Based on the deformations $\{\boldsymbol{f}_{CB}, \boldsymbol{f}_{CF}\}$ computed by the previous iteration, the discrete approach estimates the solution to Eq. (13) by first computing the intermediate warped images $\{B', F'\}$, and the intermediate posteriors $p_{B'} \triangleq p_B \circ \boldsymbol{f}_{CB}$ and $p_{F'} \triangleq p_F \circ \boldsymbol{f}_{CF}$. The deformation from the image B to the warped image B' is simply $\boldsymbol{f}_{B'B} \triangleq \boldsymbol{f}_{CB}$ (and $\boldsymbol{f}_{F'F} \triangleq \boldsymbol{f}_{CF}$ accordingly). Next, we search for the maps $\{\boldsymbol{f}_{CB'}, \boldsymbol{f}_{CF'}\}$ minimizing the energy function $E^D(\cdot,\cdot; B', F', p_{B'}, p_{F'})$, the discrete form of Eq. (13) (see Appendix B). E^D is defined on a Markov Random Field, which consists of a set of nodes $\mathcal{V}$ placed on a regular grid over the image domain Ω_C. Each node $s \in \mathcal{V}$ is associated with a pair of labels $\{l_{s,CB'}, l_{s,CF'}\}$, where the value of each label is confined to the discrete set $\mathcal{L}$. The function $\boldsymbol{d} : \mathcal{L} \rightarrow \mathbb{R}^3$ maps a label to a corresponding 3D displacement vector, e.g. $\boldsymbol{d}(l_{s,CB'})$ is the displacement of the region in Ω_C associated with node s pointing to $\Omega_{B'}$. To determine the optimal mapping $\{\boldsymbol{f}_{CB'}, \boldsymbol{f}_{CF'}\}$, we now solve the following minimization problem

$$\{l^*_{CB'}, l^*_{CF'}\} \triangleq \arg \min_{l_{CB'}, l_{CF'}} E^D(l_{CB'}, l_{CF'}; B', F', p_{B'}, p_{F'}) . \tag{20}$$

via the tree re-weighted message passing method [41], [42] (see Appendix B for further details).

Having determined the optimal labeling $l^*_{CB'}$ and $l^*_{CF'}$, we create a smooth mapping with respect to the current image resolution by computing the weighted sum of displacement vectors on a set of neighboring nodes $\mathcal{N}(\mathbf{x})$ for each voxel location $\mathbf{x} \in \Omega_C$:

$$\begin{aligned} \boldsymbol{f}_{CB'}(\mathbf{x}) &\triangleq \mathbf{x} + \sum_{s\in\mathcal{N}(\mathbf{x})} \omega_s(\mathbf{x}) \cdot \boldsymbol{d}(l^*_{s,CB'}) , \\ \boldsymbol{f}_{CF'}(\mathbf{x}) &\triangleq \mathbf{x} + \sum_{s\in\mathcal{N}(\mathbf{x})} \omega_s(\mathbf{x}) \cdot \boldsymbol{d}(l^*_{s,CF'}) . \end{aligned} \tag{21}$$

The weight $\omega_s(\mathbf{x})$ is defined by the conventional free form deformation model based on cubic B-splines [43] guaranteeing a smooth interpolation of the displacement vectors $\boldsymbol{d}(\cdot)$ defined on the grid across the entire image domain Ω_C. We note that making the interpolations of Eq. (21) exact at a coordinate $\mathbf{x}_s$ of a node $s \in \mathcal{V}$, i.e., $\boldsymbol{f}_{CB'}(\mathbf{x}_s) - \mathbf{x}_s = \boldsymbol{d}(l^*_{s,CB'})$ and $\boldsymbol{f}_{CF'}(\mathbf{x}_s) - \mathbf{x}_s = \boldsymbol{d}(l^*_{s,CF'})$, would require B-spline prefiltering [44], which we omitted for computational reasons.

Based on the mappings determined by the discrete optimization, we update the map from the original images to the center coordinate system via

$$\boldsymbol{f}_{CB} \leftarrow \boldsymbol{f}_{B'B} \circ \boldsymbol{f}_{CB'} \ , \ \ \boldsymbol{f}_{CF} \leftarrow \boldsymbol{f}_{F'F} \circ \boldsymbol{f}_{CF'} \ . \tag{22}$$

and use them to initialize the continuous optimization.

Our continuous approach determines the solution by minimizing the original, continuous energy function $E(\cdot)$ of Eq. (13) through the fluid registration scheme [45]. We iterate between

- setting the intermediate images $\{B', F'\}$, probabilities $\{p_{B'}, p_{F'}\}$, and deformations $\{\boldsymbol{f}_{B'B}, \boldsymbol{f}_{F'F}\}$
- computing the mappings from the intermediate to the center coordinate system $\{\boldsymbol{f}_{CB'}, \boldsymbol{f}_{CF'}\}$ by multiplying the gradient ∇E (as defined in Appendix C) with step size ϵ:

$$\begin{aligned} \boldsymbol{f}_{CB'} &\triangleq \mathrm{Id} - \epsilon \cdot \nabla_{f_{CB'}} E\big(\boldsymbol{f}_{CB'}, \boldsymbol{f}_{CF'}; B', F', p_{B'}, p_{F'}\big) \ , \\ \boldsymbol{f}_{CF'} &\triangleq \mathrm{Id} - \epsilon \cdot \nabla_{f_{CF'}} E\big(\boldsymbol{f}_{CB'}, \boldsymbol{f}_{CF'}; B', F', p_{B'}, p_{F'}\big) \ , \end{aligned} \tag{23}$$

- and using those results to update current mappings according to Eq. (22).

We repeat this iteration until a local optimum is found, i.e. when the gradients of Eq. (23) approach zero. After convergence, we return to the beginning of the for-loop to continue on the finer scale.

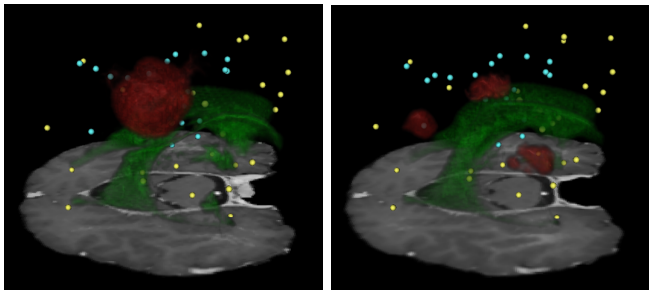
In summary, we propose a specific framework for determining the deformation between pre-operative and post-recurrence scans. Our approach, called *PORTR*, first generates an explicit model for the pathology and produces an initial mapping inferred by the tumor growth model. *PORTR* then determines the deformation between scans by confining the solution to the symmetric deformations and using the hybrid optimization method.

III. Comparative Study on 24 Subjects

We registered the pre-operative and post-recurrence scans of 24 subjects and compared *PORTR* with *DRAMMS* [20], a state-of-the-art method based on attribute vectors, mutual saliency, and discrete optimization, and *ANTS* [19], a widely used method based on symmetric registration and continuous optimization. We now first describe the experimental set up, including the data, the accuracy scores, and implementation details of each method. We then show that our method achieves the highest overall accuracy on this specific data set. We confirm the quantitative findings by visually comparing the registration results. The last experiment highlights the importance of specific components for the accuracy of our method.

A. Experimental Data

Our data set consists of 24 pairs of pre-operative and post-recurrence MR brain scans of glioma patients. We segmented each scan and had experts place landmarks in 10 pairs. We used the segmentations and landmarks to measure the accuracy of each approach. We now describe each of these components of our data set in further detail.



(a) Pre-Operative B (b) Post-Recurrence F

Fig. 8. Landmarks placed on (a) the pre-operative scan B and (b) the post-recurrence scan F (Subject 6). The landmarks of Group 1 (placed inside of 30 mm distance to the tumor boundary) are shown in cyan and the landmarks of Group 2 (placed outside of that region) are shown in yellow. Also, the tumor and cavity are shown in red and the ventricles in green. The images highlight the vast spatial distribution of the landmarks across the brain, with the landmarks of Group 1 being in close proximity to the tumor.

Each of the 24 glioma patients was scanned before surgery, referred to as pre-operative or baseline (B). The enhanced tumor region shown on the baseline scan was completely removed through surgery. The post-recurrence or follow-up scan (F) was taken after the tumor had recurred. While the time interval between the scans varied between 2 and 24 months (average 8 months), the scans themselves were acquired using the same MR acquisition protocol. Every acquisition consisted of a T1, T1-CE, T2, and FLAIR image acquired on a 3 Tesla MRI scanner systems (MAGNETOM Trio Timstem, Siemens Medical Systems, Erlangen, Germany) at the Hospital of the University of Pennsylvania. The dimension of each slice was (192×256) with pixel spacing ($0.9766 \times 0.9766\ mm^2$). T1 and T1-CE scans had 1 mm slice thickness and the T2 and FLAIR had 3 mm. Each scan was smoothed and corrected for MR field inhomogeneity [46]. Then, we co-registered T1, T2, and FLAIR to the T1-CE via affine registration based on mutual information [47]. Each modality now has the same dimension ($192 \times 256 \times 192$) and voxel size ($0.9766 \times 0.9766 \times 1.0\ mm^3$). We ended the preprocessing of the data by skull stripping [48] and affinely registering the post-recurrence to the pre-operative scans via [47].

For all 24 subjects, an expert manually segmented the tumor from the baseline scan and the cavity of the follow-up scan. We then automatically segmented the ventricles for each scan by intersecting the map inferred from the corresponding posterior of CSF (Eq. (9) or Eq. (5)) with the aligned atlas of the ventricles. The segmentations for ventricles were verified by experts.

Two experts placed landmarks on the scans of 10 randomly selected subjects. For each pre-operative scan, the first expert placed 20 landmarks inside the band defined by the 30 mm distance to the tumor boundary (Group 1) and 30 landmarks beyond the 30 mm perimeter (Group 2). The tumor boundary was inferred from the previous segmentation. The expert placed the landmarks on anatomical markers such as the bifurcations of blood vessels, the omega shape of the cortex, and midline of the brain. Both experts then independently

placed the corresponding landmarks in the post-recurrence scan. In the remainder of this section, we view the landmarks set by the first expert as the gold standard and the outcome of the second expert as a reference standard in the comparison of the automatic method.

Fig. 8 shows the landmarks placed on one pre-operative scan and the corresponding post-recurrence scan. The cyan dots represent landmarks of Group 1 and the yellow dots of Group 2. As a reference, the image also shows the tumor (red) and ventricles (green) as well as one axial T1-CE slice. The figure nicely illustrates the distribution of the landmarks, which are scattered across most of the brain area with the landmarks in Group 1 being closer to the tumor than those of Group 2.

B. Accuracy Scores

We determined the accuracy of each approach by measuring errors with respect to automatic landmark placement and overlap between aligned segmentations. The landmark error of an approach is defined as the mean distance between the landmarks aligned by the approach and the corresponding ones set by the expert. We used leave-one-out cross-validation to compute this error for all of the 10 cases with manually placed landmarks. In other words, we first determined the set of parameters of an approach that lead to the minimal overall error on 9 cases. We then recorded the landmark errors for the remaining test case by applying the method with that parameter setting to the corresponding scans. We repeated that process until we recorded the landmark errors for each of the 10 cases. After computing the average landmark error for each case and method, we then called the outcome of the two methods significantly different if the Wilcoxon signed rank test [49] between the sets of average landmark errors revealed a p-value below 0.05.

Segmentation overlap was measured across all 24 subjects. Using the previous registration results for the selected 10 landmark cases and registering the remaining 14 subjects based on the parameter setting that minimizes the landmark error across those 10 subjects, we computed the Dice score [50] between the segmentations of the aligned post-recurrence scan and the ones of the pre-operative scan. Specifically, we recorded the Dice score of the ventricle regions, which generally are severely deformed due to the mass effects of tumors, and the Dice score between the aligned cavity on the post-recurrence scan and the tumor on the pre-operative scan. Higher Dice scores indicate better registrations. We repeated the previous significance testing by replacing the landmark errors with the Dice scores.

C. Implementation Details

As previously mentioned, we compared the accuracy of *PORTR* to *ANTS* and *DRAMMS*. We now go over the specific implementations of each approach.

1) PORTR: Our method registered the scans (Step 3) by first segmenting the pathology of the follow-up scan (Step 1) and baseline scan (Step 2). In Step 1, we estimated the tumor and cavity of Eq. (2) by first finding the smallest circle that encompasses each abnormal region as shown by the hyper or hypo intensities on T1-CE. We set $a_{TU} = 0.8$ of $p_F(\mathcal{T}_{TU}|\mathbf{x})$ in Eq. (2) and $a_{ED} = 0.2$ of $p_F(\mathcal{T}_{ED}|\mathbf{x})$ in Eq. (3), which results in the slope of the radial profile of $p_F(\mathcal{T}_{TU}|\mathbf{x})$ to be steeper than that of $p_F(\mathcal{T}_{ED}|\mathbf{x})$. Furthermore, we set $b_{ED} = 4$ of $p_F(\mathcal{T}_{ED}|\mathbf{x})$ in Eq. (3) so that the radius implied by $p_F(\mathcal{T}_{ED}|\mathbf{x})$ is 4 times larger than that of $p_F(\mathcal{T}_{TU}|\mathbf{x})$. We note that our method is not very sensitive to changes in a_{TU}, a_{ED}, and b_{ED}. The variation of the mean landmark error was below 1% when varying those parameters by 20% around the chosen settings. Thus, our method is also robust towards smaller changes in the segmentation of the follow-up scans. This observation deterred us from coupling Step 1 and 2 to create a joint intra-subject registration and segmentation approach as we would expect marginal improvement at best while substantially increasing the computational burden.

Finally, we initialized $\mathbf{\Phi}_F$ in Eq. (4) by taking samples inside the corresponding tissues across all four modalities (T1, T1-CE, T2 and FLAIR). For Step 2, we repeated the previous procedures for the baselines scans estimating the initial seed location of the tumor parameter $\mathbf{q}$ in Eq. (7) and initializing $\mathbf{\Phi}_B$ of Eq. (9). The registration of Step 3 was only based on T1 and T1-CE ($N = 2$). We omitted the other two modalities (T2 and FLAIR) as their lower resolutions decrease the accuracy of the NCC measure in Eq. (14). Using both T1 and T1-CE improved the mean landmark errors by 16% compared to using T1 or T1-CE alone. The NCC measure was based on a patch width of 9 voxels. In Eq. (18), we used $c_i = \sigma^{2i}/(i! \cdot 2^i)$ so that we can minimize the smoothness term E_S by applying the Gaussian kernel with standard deviation σ to the gradients of E_C and E_P [36] (see Appendix C). We fixed $\sigma = \sqrt{3}$ in all experiments. We then determined the optimal weighing parameters λ_D and λ_P of Eq. (13) via leave-one-out cross-validation. The search space of λ_D was $[0.8, 1.2]$ and of λ_P was $[0.1, 0.3]$. This implementation of *PORTR* is freely available for download via the website of the Section of Biomedical Image Analysis, University of Pennsylvania.

2) DRAMMS and mDRAMMS: We choose *DRAMMS* [20] as a representative of registration methods based on discrete optimization. Its mutual-saliency concept is well suited for our data set, which requires the registration of scans with missing correspondences. *DRAMMS* produced the best results based just on the T1 modality. Note that *DRAMMS* currently works only for a single-channel. Thus, it cannot take advantage of the multiple channels such as the other methods of this comparison and might therefore be at a disadvantage in our comparisons. We also included a second implementation of *DRAMMS* in our comparison, called *mDRAMMS*, which is guided by the segmentation of pathology for pre-operative scans generated by our approach in Step 2. Specifically, we confined *mDRAMMS* to the mask of the baseline scan defined by the complement of the posterior of pathological regions $1 - p_{B,PT}$ of Eq. (15). The mask for the follow-up scan was omitted as the current publicly available version of *DRAMMS* does not accept it as input. For each implementation, we searched for the optimal regularization parameter g in the range of $[0.1, 0.5]$ and mutual saliency parameter c by setting it to 0 or 1. These ranges were suggested by the creators of *DRAMMS* for registering image pairs with large deformation.

TABLE I
AVERAGE RUNNING TIME

PORTR	*DRAMMS*	*mDRAMMS*	*ANTS*	*mANTS*
3.5 h	0.8 h	0.7 h	1.7 h	1.2h

3) ANTS and mANTS: We choose *ANTS* [19] as a representative for registration methods based on continuous optimization. Its symmetric registration scheme is well suited for the large mass effects caused by the tumor. In addition, the method compares favorably to other approaches in various registration tasks [51], [52]. For the same reason as with *PORTR*, we achieved the highest accuracy confining *ANTS* to T1 and T1-CE channels. Like *mDRAMMS*, we also include a second implementation in our comparison, called *mANTS*, which used $1 - p_{B,PT}$ as a mask of the baseline scan and ignored the mask for the follow-up scan. The publicly available version of *ANTS* currently cannot be constrained by the mask of the follow-up scan. Each implementation used cross correlation (CC) to measure image similarity, hierarchically iterated based on $(100 \times 100 \times 50)$, and used the symmetric image normalization (SyN) scheme. We furthermore determined the optimal setting during cross validation for the step-size s of the SyN scheme in the range of $[0.25, 0.5]$ and the regularization on the deformation field t in the range of $[0.0, 1.5]$, where $t = 0.0$ allows maximum flexibility. These intervals were chosen based on the recent evaluation by the creators of *ANTS* [53].

We end the description of our applications by mentioning their running time summarized in Table I. On an Intel Core i7 3.4 GHz machine with Windows operating system, *PORTR* average running time was 3.5 h (less than ten minutes for Step 1, 1.5 h for Step 2, and 1.9 h for Step 3) while *ANTS* took 1.7 h and *mANTS* 1.2 h. *DRAMMS* and *mDRAMMS* took 0.8 and 0.7 hours respectively on an Intel Xeon 3.06 GHz machine with Linux operating system.

D. Registration Results

We now compare the accuracy of each implementation on our data set of 24 subjects in three steps. We first review the landmark-based error followed by the segmentation-based error. Then we visually compare the results, which confirm the findings of the two quantitative evaluations. We end with checking the role of specific components of *PORTR*. Note that the baseline for comparison is the outcome of affine registration [47] referred to as *AFFINE*, and those of the second rater referred to as *RATER*.

1) Landmark-Based Errors: Fig. 9 shows the box-and-whisker plots of average landmark errors based on landmarks inside the 30 mm tumor boundary (Group 1) in the top graph as well as the one for the remaining landmarks (Group 2) in the bottom graph across the 10 subjects. The error statistics were computed with respect to the distance of the landmarks set by the first rater. For each method, the black dot represents the mean landmark error.

The errors of all non-rigid registration methods are significantly lower than those of *AFFINE*. Among non-rigid registration methods, *PORTR* has the lowest mean error, closest to that of *RATER*. The mean error of *PORTR* is 25% lower than *DRAMMS*, 24% lower than *mDRAMMS*, 9% lower than *ANTS*, and 7% lower than *mANTS* for landmarks of Group 2. For landmarks nearby tumor (Group 1), these performance gaps respectively widen to 46%, 42%, 38%, and 34%. *PORTR* was significantly better than the other competing methods with respect to the landmark error of Group 1 ($p < 0.01$) as well as Group 2 ($p < 0.05$).

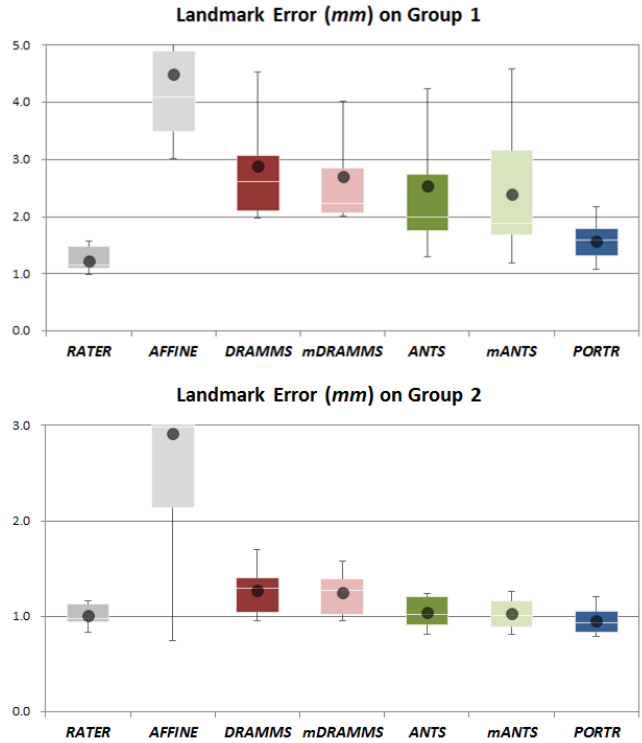


Fig. 9. The box-and-whisker plots of average landmark errors evaluated using landmarks of Group 1 (top) and landmarks of Group 2 (bottom) across the 10 subjects. The bars start at the lower quartile and end at the upper quartile with the white line representing the median. The whiskers show minimum and maximum values within 1.5 interquartile ranges from lower and upper quartile, respectively (outliers are not shown). The black dots represent the mean landmark errors. *RATER* denotes the landmark errors of the second rater and *AFFINE* shows the errors of the affine registration. Among the registration approaches, *PORTR* performs best and has the lowest mean score and smallest variation.

Among the alternative methods, *ANTS* performed better than *DRAMMS*, however the difference between the mean errors is smaller than that between *PORTR* and *ANTS*. The methods with tumor masks (*mDRAMMS* or *mANTS*) performed similar (performance gaps are less than 5%) to their counterparts (*DRAMMS* or *ANTS*) as these methods assume smooth deformations inside the masked tumor regions. This assumption is inaccurate with respect to recovering mass effects.

We note that landmarks were placed in regions that could be clearly recognized anatomically by the experts. As many tumors induce large deformations and great signal changes around them, identifying such landmarks very close to the tumor is nearly impossible. Therefore, it is likely that the true registration error in the immediate vicinity of the tumor is larger than the error measured in Fig. 9.

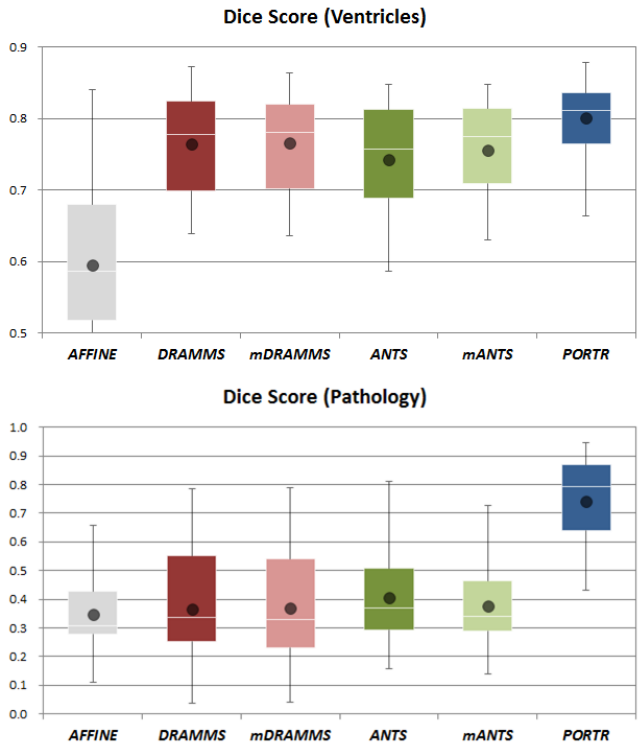


Fig. 10. The box-and-whisker plots of Dice scores evaluated on segmentations of ventricles (top) and pathology (bottom) across the 24 subjects. The results show that *PORTR* performs better than the other approaches for ventricles and pathology.

2) Segmentation-Based Errors: Fig. 10 summarizes the Dice score of each implementation across the 24 subjects. For the ventricles (top graph) and the pathology (bottom graph), *AFFINE* received the lowest mean score. As expected, *AFFINE* performed worst as the registration does not have enough degrees of freedom to model the impact of pathology on all brain structures. All other methods were fairly accurate in registering the ventricles. We note that none of the approaches, including *PORTR*, explicitly modeled this anatomy in their cost function. Thus, the Dice score of the ventricles provides an unbiased comparison across the methods. In this comparison, the mean score of *PORTR* is at least 3% better than that of any other method. Overall, *PORTR* was significantly better than the other competing methods ($p < 0.0001$).

Fig. 10 (bottom graph) shows the Dice scores for the pathological regions. With the exception of *PORTR*, the scores of the automatic methods significantly dropped compared to the scores achieved for the ventricles. Out of those methods, *ANTS* performed slightly better with a mean Dice score of 41%. Interestingly, the implementations based on the tumor masks (*mDRAMMS* or *mANTS*) did not perform better than their counterparts (*DRAMMS* or *ANTS*). This indicates that simply masking tumor regions does not lead to better overlaps on pathological regions. Unlike the other methods, *PORTR* explicitly matched the pathologies across scans via Eq. (17). The explicit modeling enabled our approach to achieve quite good accuracy with an average score of 74%, which is 33% better than *ANTS*. Overall, *PORTR* was significantly better than the other competing methods ($p < 0.0001$).

3) Visual Comparisons: We now visually compare the registration results of 10 subjects used for measuring landmark errors. Fig. 11 (a) shows the T1-CE image of the baseline scan with the tumor outlined in red and ventricles in green. Fig. 11 (b) shows the corresponding follow-up scan. Fig. 11 (c)-(g) show the follow-up scan registered to the baseline according to each method. As a reference, the tumor (red) and ventricles (green) of the baseline are overlaid in the aligned scans.

The images confirm our quantitative findings. For each subject, the aligned follow-up of *PORTR* much better matches the baseline scan than those of other competing methods. The ventricles of the follow-up scans aligned by *PORTR* overlap well with the baseline across all examples. This is not the case for the results of the competing methods where the ventricles leak to the adjacent tumor regions in Subjects 3 and 6. Furthermore, the ventricles inaccurately match in Subjects 2, 7 and 10. In Subject 4, 5 and 9, all methods align the ventricle regions well as the tumor is distant from ventricles. For the registration quality around pathology, *PORTR* well aligns tumor and cavity regions in all examples. However, results of other competing methods generally failed to produce reasonable overlaps on pathological regions except Subject 9 where the mass effect is small.

Interestingly, there are no big visual differences between *DRAMMS* and *mDRAMMS*. We presume that their mutual-saliency term puts low confidences on pathological regions, so the tumor masks do not greatly help in those regions. On the other hand, the results of *mANTS* look different from those of *ANTS*, especially on Subject 1, 3, and 7, but not necessarily improved. *mANTS* tends to preserve the appearance of follow-up scans in pathological regions as the region is masked out in the corresponding energy function. For *mDRAMMS* and *mANTS*, the tumor masks only assist in maintaining smooth deformations on pathological regions. The poor matches by the four competing methods (*ANTS*, *mANTS*, *DRAMMS* and *mDRAMMS*) on pathological regions thus indicates that it is hard to match pathological regions between baseline and follow-up scans using imaging information alone.

Next, we review the quality of our registration specifically in cortical regions nearby the tumor. We do so in Fig. 12 by taking a closer look at two examples: the registration results with respect to Subjects 5 and 7. The red arrow in Subject 5 points to the cortex, whose shape in the aligned image by *PORTR* (g) matches the one in the original image (a). This is not the case for the results generated by the other methods (c)-(f). *PORTR* is also the only method where the cortical region around the recurrence (yellow arrow) is properly aligned to the baseline scan. It does so by dramatically shrinking the recurrence in the aligned scans, which the other methods failed to do. In Subject 7, *PORTR* is again the only method that accurately aligns the cortex region pointed out by the red arrow. While the other methods try to match tumor recurrence to the original tumor, our approach correctly aligns the resection cavity to the pathology. Especially in this case, the goal of *PORTR* to match the resection cavity to the tumor seems to help in registering the healthy tissue.

In summary, *PORTR* produced the visually the most rea-

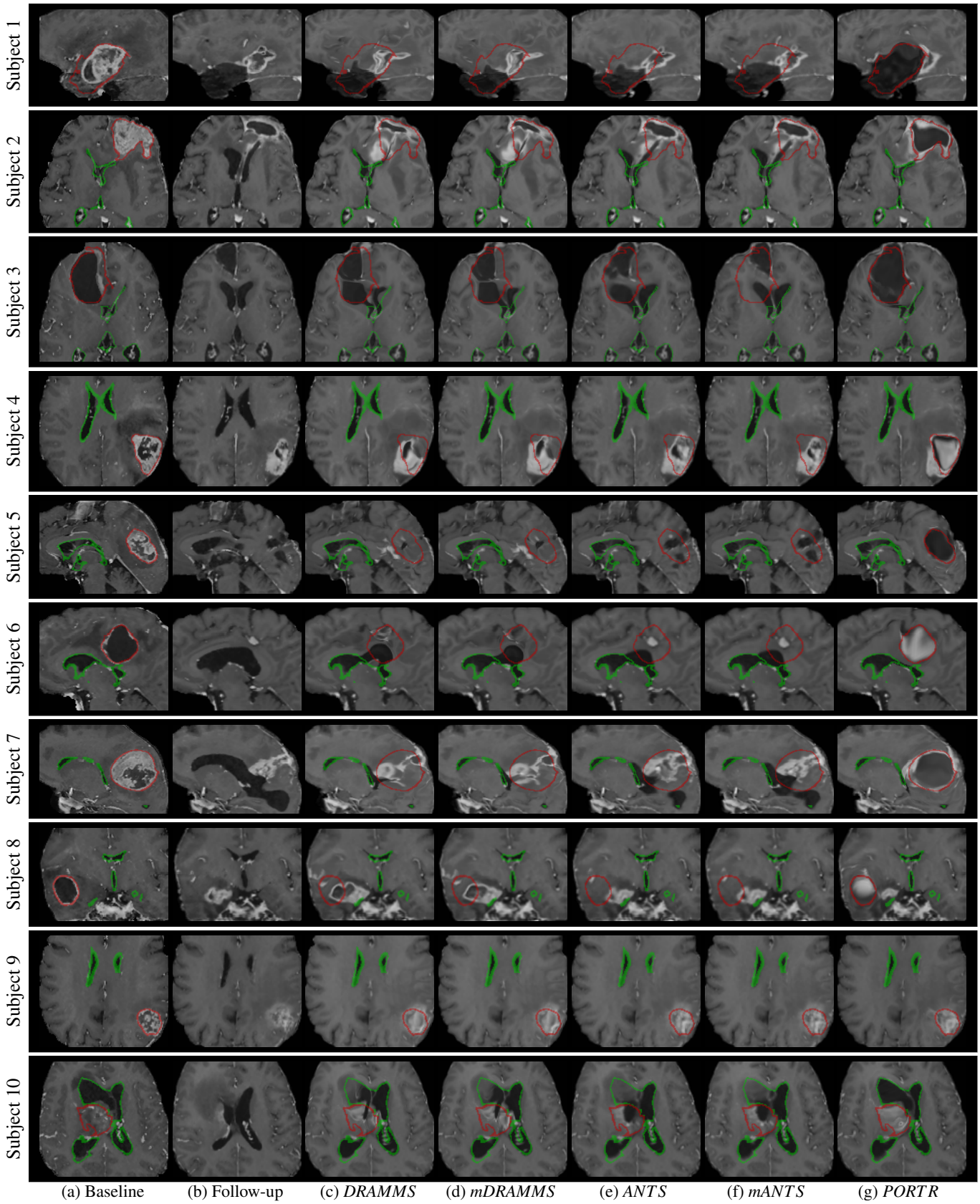


Fig. 11. The registration results of follow-up onto baseline scans. In each row, we show T1-CE images of the pre-operative scan (baseline) in (a) and the post-recurrence scan (follow-up) in (b). (c)-(g) show the registered post-recurrence scans using *DRAMMS*, *mDRAMMS*, *ANTS*, *mANTS*, and *PORTR*, respectively. For baseline and registered scans, boundaries of segmented tumor (red) and ventricles (green) of baseline are overlaid. Based on visual comparison of these images, *PORTR* shows more reasonable results than the other non-rigid registration methods in all 10 cases.

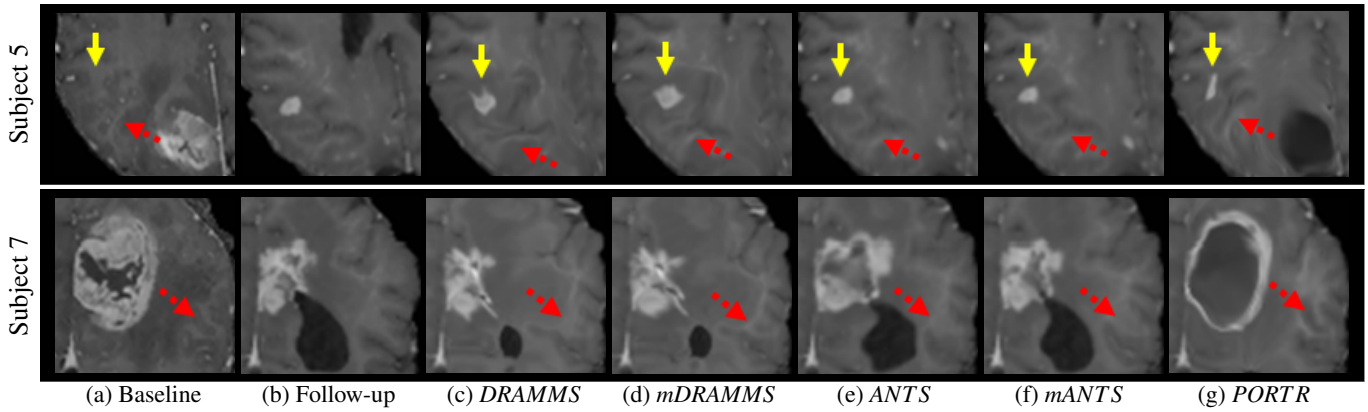


Fig. 12. The magnified views of follow-up onto baseline scans for selected subjects. The figures are shown in the same order as in Fig. 11. In each figure, the red arrow marks cortical structure and the yellow arrow marks tumor recurrence. Based on visual comparisons, The results of *PORTR* outperforms the other methods.

sonable results among the non-rigid registration methods. In all cases, the resection cavity of the follow-up scan properly overlapped with the tumor on the baseline scan. The same is true for the ventricles. Overall, the visual results echoed our quantitative findings based on landmark end segmentation error.

E. Role of Specific Components of PORTR

We follow up the previous comparisons by taking a closer look at the different components of *PORTR*. Specifically, we analyze the role of the hybrid optimization method, the symmetric framework, the pathology term, and the initial mapping in Step 3.

We further validate our chosen registration framework by confining *PORTR* to the discrete optimization (called *Discrete*), the continuous optimization (called *Continuous*), and by replacing the symmetric approach with directly mapping B to F (called *Asymm*). In *Asymm*, $\boldsymbol{f}_{CB}$ is fixed to the identity so that $\boldsymbol{f}_{CF}$ is actually the mapping $\boldsymbol{f}_{BF}$. The landmark errors for these three implementations are summarized in Fig. 13. As expected, *PORTR* produced lower errors than the other methods with respect to the landmarks of Group 1 as well as Group 2. For landmarks nearby tumor (Group 1), the mean error of *PORTR* is 20% lower than *Discrete*, 7% lower than *Continuous*, and 10% lower than *Asymm*. Overall, the error scores of *PORTR* were significantly lower than those of *Discrete* and *Asymm* for Group 1 ($p < 0.01$) and that of *Discrete* for Group 2 ($p = 0.0120$). Compared to *Continuous*, *PORTR* may be better with respect to Group 2 ($p = 0.0969$) but proving this hypothesis would require additional error measurements. Combining these results, *PORTR* improved the performances compared to its simplified versions, which further justifies our design choices.

Next, we analyze the impact of the results generated in Step 1 and 2 on the accuracy of *PORTR*. We first ran *PORTR* with λ_p in Eq. (13) set to zero (called *Without* E_P). In other words, *Without* E_P ignored the tumor matching term E_P during registration. We also ran *PORTR* by setting the initial mapping $\boldsymbol{f}^0_{BF}$ of Eq. (10) to the identity function (called *Without* $\boldsymbol{f}^0_{BF}$). Thus, *Without* $\boldsymbol{f}^0_{BF}$ ignored the deformation computed in Step 2. Fig. 14 summarizes the segmentation-based error of both implementations on the 24 subjects. With respect to pathology, the mean Dice value of *PORTR* is 13% higher than *Without* E_P and 6% higher than *Without* $\boldsymbol{f}^0_{BF}$. In comparison, the differences of mean scores

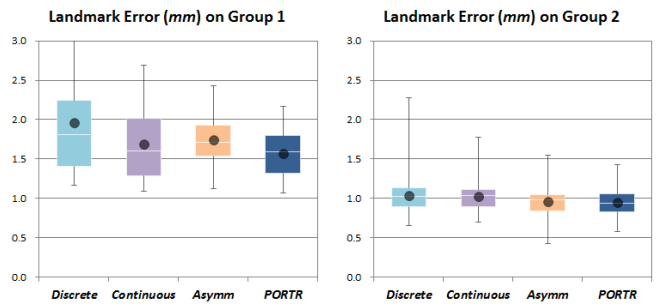


Fig. 13. The box-and-whisker plots of average landmark errors with respect to Group 1 (left) and Group 2 (right) obtained by changing the optimization of *PORTR*. Specifically, the graphs compare the accuracy of *Discrete* (*PORTR* running the discrete optimization part only), *Continuous* (*PORTR* running the continuous optimization part only), *Asymm* (the asymmetric version of *PORTR*), and *PORTR*. The results indicate that *PORTR* performs better than the other variants.

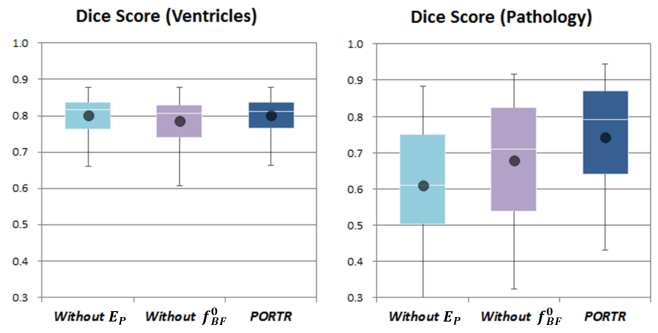


Fig. 14. The box-and-whisker plots of Dice scores with respect to the segmentations of ventricles (left) and pathology (right). The implementations listed on the horizontal axis are *Without* E_P (*PORTR* without the term E_P in Eq. (17)), *Without* $\boldsymbol{f}^0_{BF}$ (*PORTR* initialized with the identity function), and *PORTR*. The results show *PORTR* performs better than *Without* E_P and *Without* $\boldsymbol{f}^0_{BF}$ for pathological regions while they have similar Dice scores with respect to the ventricles.

are less than 2% for the ventricles. Overall, the Dice scores of *PORTR* with respect to pathology were significantly better than those of *Without* E_P and *Without* $\boldsymbol{f}^0_{BF}$ ($p < 0.0001$), which further motivates the need for the information gained from Step 1 and 2 of our registration framework.

In summary, our experiments show that the proposed method is more accurate for the registration of pre-operative and post-recurrence glioma scans than certain state-of-the art approaches. Our method achieved the highest accuracy in the landmark comparison, produced the most plausible deformations on pathological regions, and received the highest Dice scores with respect to ventricles and pathologies.

IV. Conclusion

We presented a new deformable registration approach that matches intensities of healthy tissue as well as glioma to resection cavity. Our method extracted pathological information on both scans using scan-specific approaches and then registers scans by combining image-based matchings with pathological information. To achieve unbiased deformation fields on either scan, we used a symmetric formulation of our energy model comprised of image- and shape-based correspondences and smoothness constraints. We determined the optimal registration results by minimizing the energy function using a hybrid optimization strategy which takes advantages both of discrete and continuous optimizations. We compared our approach to state-of-the-art registration methods in registering pre-operative and post-recurrence MR scans of 24 glioma patients. We quantitatively compared their outcome with respect to matching landmarks and segmentations, following up this comparison with visual inspection. In this comparison, our approach performed significantly better than the other registration methods.

Acknowledgment

The authors would like to thank Dr. Y. Ou for helping us to run *DRAMMS* and Dr. H.S. Javitz for his advice on the statistical analysis.

Appendix A
A Bayesian Model for Joint Segmentation and Registration

We now describe in detail our approach for joint segmentation and registration in Step 1. As defined in Step 1, $\boldsymbol{h}_F$ is the unknown vector field representing the mapping from Ω_F to the atlas space and $\boldsymbol{\Phi}_F$ is the unknown intensity distributions of the different tissue classes. Inspired by Ashburner and Friston [54] and Pohl *et al.* [55], [56], one way to jointly compute the probabilities and align the atlas is by solving the following maximum a posteriori (MAP) estimation problem:

$$\begin{aligned}\{\boldsymbol{h}^*_F, \boldsymbol{\Phi}^*_F\} &= \arg\max_{\boldsymbol{h}_F, \boldsymbol{\Phi}_F} \; p_F(\boldsymbol{h}_F, \boldsymbol{\Phi}_F | F) \\ &= \arg\max_{\boldsymbol{h}_F, \boldsymbol{\Phi}_F} \left(\sum_{\mathcal{T}} p_F(\boldsymbol{h}_F, \boldsymbol{\Phi}_F, \mathcal{T} | F) \right) , \end{aligned} \quad (24)$$

where we marginalize over $\mathcal{T}$ to simplify the modeling.

To decompose this MAP problem, we make use of the following independence assumptions: F is independent of $\boldsymbol{h}_F$ conditioned $\mathcal{T}$, $\mathcal{T}$ is independent of $\boldsymbol{\Phi}_F$ conditioned $\boldsymbol{h}_F$, $\boldsymbol{h}_F$ is independent of $\boldsymbol{\Phi}_F$, and $\mathcal{T}$ is composed as a set of independent random variables across the image grid Ω_F. $p_F(\mathcal{T}|\boldsymbol{h}_F)$ and likelihoods $p_F(F|\mathcal{T}, \boldsymbol{\Phi}_F)$ are defined by the product of the corresponding probabilities over all the voxels in Ω_F. Then Eq. (24) simplifies to:

$$\begin{aligned}\{\boldsymbol{h}^*_F, \boldsymbol{\Phi}^*_F\} &= \arg\max_{\boldsymbol{h}_F, \boldsymbol{\Phi}_F} \left(\sum_{\mathcal{T}} p_F(\mathcal{T}|\boldsymbol{h}_F) \cdot p_F(F|\mathcal{T}, \boldsymbol{\Phi}_F) \right) \\ &= \arg\max_{\boldsymbol{h}_F, \boldsymbol{\Phi}_F} \left(\prod_{\mathbf{x} \in \Omega_F} \sum_{t \in \Theta} p_F(\mathcal{T}_t|\boldsymbol{h}_F, \mathbf{x}) \cdot p_F(F|\mathcal{T}_t, \boldsymbol{\Phi}_F, \mathbf{x}) \right) . \end{aligned} \quad (25)$$

Note that we dropped terms not depending on $\boldsymbol{h}_F$ or $\boldsymbol{\Phi}_F$.

We define the first term of the above equation, $p_F(\mathcal{T}_t|\boldsymbol{h}_F, \mathbf{x})$, through deforming our atlas $p_F(\mathcal{T}_t|\mathbf{x})$ via $\boldsymbol{h}_F$:

$$p_F(\mathcal{T}_t|\boldsymbol{h}_F, \mathbf{x}) \triangleq p_F(\mathcal{T}_t|\boldsymbol{h}_F(\mathbf{x})) \; . \quad (26)$$

We model the second term, the image likelihood $p_F(F|\mathcal{T}_t, \boldsymbol{\Phi}_F, \mathbf{x})$, as a multivariate Gaussian with the tissue specific mean $\mathbf{m}_t$ and covariance $\boldsymbol{\Sigma}_t$ composing $\boldsymbol{\Phi}_F$. We obtain Eq. (4) by applying Eq. (26) on Eq. (25).

Ashburner and Friston [54] and Pohl *et al.* [55], [56] have shown that the solutions to problems such as Eq. (25) can robustly be estimated via the EM algorithm [28]. The EM algorithm iteratively determines the solution by computing the posterior

$$\begin{aligned} p_F(\mathcal{T}_t|F, \mathbf{x}) &\triangleq p_F(\mathcal{T}_t|F, \boldsymbol{h}'_F, \boldsymbol{\Phi}'_F, \mathbf{x}) \\ &\propto p_F\big(\mathcal{T}_t|\boldsymbol{h}'_F(\mathbf{x})\big) \cdot p_F(F|\mathcal{T}_t, \boldsymbol{\Phi}'_F, \mathbf{x}) \end{aligned}$$

in the E-Step and updating in the M-Step the parameters

$$\boldsymbol{\Phi}'_F \leftarrow \arg\max_{\boldsymbol{\Phi}_F} \left(\sum_{\mathbf{x} \in \Omega_F} \sum_{t \in \Theta} p_F(\mathcal{T}_t|F, \mathbf{x}) \cdot \log p_F(F|\mathcal{T}_t, \boldsymbol{\Phi}_F, \mathbf{x}) \right)$$

which is solved in a closed form of [57], and

$$\boldsymbol{h}'_F \leftarrow \arg\max_{\boldsymbol{h}_F} \left(\sum_{\mathbf{x} \in \Omega_F} \sum_{t \in \Theta} p_F(\mathcal{T}_t|F, \mathbf{x}) \cdot \log p_F(\mathcal{T}_t|\boldsymbol{h}_F, \mathbf{x}) \right)$$

which iteratively can be solved as in [16]. After convergence, we assign $\boldsymbol{h}'_F$ and $\boldsymbol{\Phi}'_F$ to $\boldsymbol{h}^*_F$ and $\boldsymbol{\Phi}^*_F$, respectively.

Appendix B
Energy Functions for Discrete Optimization

We now specify the discrete version of our energy model in Eq. (13) based on the input $\{B', F', p_{B'}, p_{F'}\}$. This discrete version is based on a Markov Random Field (MRF) model that consists of a set of nodes $\mathcal{V}$ placed on a cubic grid in Ω_C and a set of hyperedges $\mathcal{E}$, where each edge is defined by three successive nodes on one axis [58], [59]. For example on the x-axis (and y-axis and z-axis accordingly), one hyperedge is defined for each set of nodes $(x-1, y, z)$, (x, y, z), and $(x+1, y, z)$. We restrict the maximum displacement of the discrete optimization to 0.4 times of the spacing between neighboring nodes ensuring that the resulting deformation is diffeomorphic [60]. Using the notations in Step 3, we define

the correspondence term of Eq. (15) as

$$E_C^D(l_{CB'}, l_{CF'}; B', F', p_{B'}, p_{F'}) \approx \sum_{s\in\mathcal{V}} \Big\{1-p_{B',PT}\big(\mathbf{x}_s+\boldsymbol{d}(l_{s,CB'})\big)\Big\}\cdot\Big\{1-p_{F',PT}\big(\mathbf{x}_s+\boldsymbol{d}(l_{s,CF'})\big)\Big\} \cdot D_{NCC}\big(B', F', \mathbf{x}_s + \boldsymbol{d}(l_{s,CB'}), \mathbf{x}_s + \boldsymbol{d}(l_{s,CF'})\big)\big) \ , \quad (27)$$

where $\mathbf{x}_s$ is a coordinate of a node $s \in \mathcal{V}$. For D_{NCC}, we use slightly different definition of Eq. (14). Let us define the region of the patch on I as $\mathcal{R}(\mathbf{x}_I)$ centered on $\mathbf{x}_I$ for each $I \in \{B', F'\}$. Then the NCC between two patches respectively centered on $\mathbf{x}_{B'}$ and $\mathbf{x}_{F'}$ is defined as

$$D_{NCC}\big(B', F', \mathbf{x}_{B'}, \mathbf{x}_{F'}\big) \triangleq \frac{1}{N}\sum_{i=1}^{N} \frac{\langle \overline{B'_i}, \overline{F'_i}\rangle}{\sqrt{\langle \overline{B'_i}, \overline{B'_i}\rangle \cdot \langle \overline{F'_i}, \overline{F'_i}\rangle}} \ ,$$

where $m(\mathbf{x}_I)$ is the mean value of the patch and

$$\overline{I_i} = \{I_i(\mathbf{y}) - m(\mathbf{x}_I) \mid \mathbf{y} \in \mathcal{R}(\mathbf{x}_I)\} \quad (28)$$

is an intensity corrected patch for each $I \in \{B', F'\}$. As we measure NCC between translated patches, this function approximates Eq. (15). For discrete optimizations, it is currently intractable to solve the exact conversion of Eq. (15) as it introduces higher-order potentials encoding each movement of the neighboring nodes. Next, we discretize the pathology term of Eq. (17) as follows

$$E_P^D(l_{CB'}, l_{CF'}; p_{B'}, p_{F'}) = \sum_{s\in\mathcal{V}} \Big\{p_{B',TU}\big(\mathbf{x}_s+\boldsymbol{d}(l_{s,CB'})\big)-p_{F',TU}\big(\mathbf{x}_s+\boldsymbol{d}(l_{s,CF'})\big)\Big\}^2 \ . \quad (29)$$

Finally we convert the smoothness term in Eq. (19) as follows:

$$E_S^D(l_{CB'}, l_{CF'}) \approx \sum_{(s,t,u)\in\mathcal{E}} \Big\{\big|\big|\boldsymbol{d}(l_{s,CB'}) - 2\cdot\boldsymbol{d}(l_{t,CB'}) + \boldsymbol{d}(l_{u,CB'})\big|\big|^2 + \big|\big|\boldsymbol{d}(l_{s,CF'}) - 2\cdot\boldsymbol{d}(l_{t,CF'}) + \boldsymbol{d}(l_{u,CF'})\big|\big|^2\Big\} \ , \quad (30)$$

where $||\cdot||$ is ℓ^2-norm. We incorporate a second-order smoothness prior [58], [59] as an approximation of the regularization in Eq. (19). The second order prior is selected as it produces smoother deformations than the first order one conventionally used in discrete registration approaches [18]. The discrete energy function E^D is defined as a weighted sum of Eqs. (27)-(30) using λ_D and λ_P as in Eq. (13):

$$E^D\big(l_{CB'}, l_{CF'}; B', F', p_{B'}, p_{F'}\big) \triangleq \lambda_D \cdot E_C^D\big(l_{CB'}, l_{CF'}; B', F', p_{B'}, p_{F'}\big) + \lambda_D \cdot \lambda_P \cdot E_P^D\big(l_{CB'}, l_{CF'}; p_{B'}, p_{F'}\big) + E_S^D\big(l_{CB'}, l_{CF'}\big) \ . \quad (31)$$

According to Eq. (20), we obtain $\{l^*_{CB'}, l^*_{CF'}\}$ by determining the label minimizing Eq. (31). However, this task is difficult as the complexity of the solution space is in $\mathcal{O}(\mathcal{L}^2)$. Instead, we perform coordinate descent:

$$l^*_{CB'} \leftarrow \arg\min_{l_{CB'}} \ E^D(l_{CB'}, l^*_{CF'}; B', F', p_{B'}, p_{F'}) \ , \quad (32)$$

$$l^*_{CF'} \leftarrow \arg\min_{l_{CF'}} \ E^D(l^*_{CB'}, l_{CF'}; B', F', p_{B'}, p_{F'}) \ . \quad (33)$$

We initialize each label with the zero displacement $\boldsymbol{d}(l^*_{s,CB'}) = \boldsymbol{d}(l^*_{s,CF'}) = 0$ and repeat solving the two minimization problems until the labels converge.

We solve Eq. (32) (and Eq. (33) accordingly) taking advantage of the fact that $l^* = l^*_{CF'}$ is fixed so that we can reduce $E^D(\cdot)$ to the parts that depend on $l = l_{CB'}$ and omit all others, i.e.

$$E'^D(l) = \sum_{s\in\mathcal{V}} \theta_s(l_s) + \sum_{(s,t,u)\in\mathcal{E}} \theta_{stu}(l_s, l_t, l_u) \quad (34)$$

with the unary potential

$$\theta_s(l_s) \triangleq \lambda_D \cdot \{1 - p_{B',PT}(\mathbf{x}_s + \boldsymbol{d}(l_s))\} \cdot \{1 - p_{F',PT}(\mathbf{x}_s + \boldsymbol{d}(l^*_s))\} \cdot D_{NCC}(B', F', \mathbf{x}_s + \boldsymbol{d}(l_s), \mathbf{x}_s + \boldsymbol{d}(l^*_s))) + \lambda_D \cdot \lambda_P \cdot \{p_{B',TU}(\mathbf{x}_s + \boldsymbol{d}(l_s)) - p_{F',TU}(\mathbf{x}_s + \boldsymbol{d}(l^*_s))\}^2$$

defined according to Eq. (27) + Eq. (29), and the ternary potential

$$\theta_{stu}(l_s, l_t, l_u) \triangleq ||\boldsymbol{d}(l_s) - 2\cdot\boldsymbol{d}(l_t) + \boldsymbol{d}(l_u)||^2$$

defined according to Eq. (30). The solution of E'^D is the same as that of Eq. (32), which we determine via the tree re-weighted message passing method (TRW) [41], [42]. We choose TRW as it performed favorably in comparison to the state-of-the-art on related discrete optimization tasks [61].

As TRW works only on pairwise MRFs, we convert θ_{stu} into pairwise potentials by creating for each edge $(s,t,u) \in \mathcal{E}$ an auxiliary node α. The node α takes on label $z_\alpha \in \mathcal{Z}$, where $\mathcal{Z}$ is a combination of the label spaces defined for s, t, and u. We assume any value of z_α has one-to-one correspondence with a triplet (z_s, z_t, z_u) where $\{z_s, z_t, z_u\} \in \mathcal{L}$. We now define a pairwise potential $\psi_{\alpha i}(\cdot)$ penalizing inconsistencies between the auxiliary node α and the (ordinary) node $i \in \{s,t,u\}$ as

$$\psi_{\alpha i}(z_\alpha, l_i) = \begin{cases} 0 & \text{if } z_i = l_i \\ \infty & \text{otherwise} \end{cases} \ ,$$

and the unary, data potential $\psi_\alpha(\cdot)$

$$\psi_\alpha(z) = \theta_{stu}(z_s, z_t, z_u) \ ,$$

so that

$$\theta_{stu}(l_s, l_t, l_u) = \min_{z_\alpha}\Big\{\psi_\alpha(z_\alpha) + \sum_{i\in\{s,t,u\}} \psi_{\alpha i}(z_\alpha, l_i)\Big\} \ . \quad (35)$$

Let $\mathcal{V}_\mathcal{A}$ be a set of auxiliary nodes and $\mathcal{E}_\mathcal{A}$ be a set of edges between auxiliary nodes $\alpha \in \mathcal{V}_\mathcal{A}$ and ordinary nodes $i \in \mathcal{V}$. Using Eq. (35), we convert the energy function of Eq. (34) into an energy function of an MRF model with

pairwise potentials

$$E''^D(l,z) = \sum_{s\in\mathcal{V}} \theta_s(l_s) + \sum_{\alpha\in\mathcal{V}_\mathcal{A}} \psi_\alpha(z_\alpha) + \sum_{(\alpha,i)\in\mathcal{E}_\mathcal{A}} \psi_{\alpha i}(z_\alpha, l_i) \ ,$$

which we then plug into TRW to determine the solution $\{l^*, z^*\} = \arg\min_{l,z} E''^D(l,z)$. Note that l^* minimizes Eq. (32) as $\min_z E''^D(l,z) = E'^D(l)$.

Appendix C
Gradients for Continuous Optimization

We now determine the gradients for our continuous optimization of Step 3 based on the input $\{B', F', p_{B'}, p_{F'}\}$. The gradients in Eq. (23) are defined as

$$\begin{aligned} \nabla_{f_{CB'}} E &= G_\sigma * \left(\lambda_D \cdot \nabla_{f_{CB'}} E_C + \lambda_D \cdot \lambda_P \cdot \nabla_{f_{CB'}} E_P\right) \ , \\ \nabla_{f_{CF'}} E &= G_\sigma * \left(\lambda_D \cdot \nabla_{f_{CF'}} E_C + \lambda_D \cdot \lambda_P \cdot \nabla_{f_{CF'}} E_P\right) \ , \end{aligned} \tag{36}$$

where G_σ is the Gaussian kernel with standard deviation σ, which is the Green's function for the Tikhonov regularization of Eq. (18) with $c_i = \sigma^{2i}/(i! \cdot 2^i)$ [36]. The gradients for the correspondence term in Eq. (15) are defined as

$$\begin{aligned} \nabla_{f_{CB'}} E_C(\mathbf{x}) &= \left(1 - p_{B',PT}(\mathbf{x})\right) \cdot \left(1 - p_{F',PT}(\mathbf{x})\right) \\ &\quad \cdot \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\sqrt{\nu_i \cdot \omega_i}} \left(\overline{F'_i}(\mathbf{x}) - \frac{\mu_i}{\nu_i} \cdot \overline{B'_i}(\mathbf{x})\right) \cdot \nabla \overline{B'_i}(\mathbf{x}) \ , \\ \nabla_{f_{CF'}} E_C(\mathbf{x}) &= \left(1 - p_{B',PT}(\mathbf{x})\right) \cdot \left(1 - p_{F',PT}(\mathbf{x})\right) \\ &\quad \cdot \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\sqrt{\nu_i \cdot \omega_i}} \left(\overline{B'_i}(\mathbf{x}) - \frac{\mu_i}{\omega_i} \cdot \overline{F'_i}(\mathbf{x})\right) \cdot \nabla \overline{F'_i}(\mathbf{x}) \ , \end{aligned} \tag{37}$$

where $\overline{B'_i}$ and $\overline{F'_i}$ are intensity corrected patches defined in Eq. (28) and we set $\mu_i \triangleq \langle \overline{B'_i}, \overline{F'_i} \rangle$, $\nu_i \triangleq \langle \overline{B'_i}, \overline{B'_i} \rangle$, and $\omega_i \triangleq \langle \overline{F'_i}, \overline{F'_i} \rangle$. The gradients for the pathology term in Eq. (17) are

$$\begin{aligned} \nabla_{f_{CB'}} E_P(\mathbf{x}) &= 2 \cdot \left(p_{B',TU}(\mathbf{x}) - p_{F',TU}(\mathbf{x})\right) \cdot \nabla p_{B',TU}(\mathbf{x}) \ , \\ \nabla_{f_{CF'}} E_P(\mathbf{x}) &= 2 \cdot \left(p_{F',TU}(\mathbf{x}) - p_{B',TU}(\mathbf{x})\right) \cdot \nabla p_{F',TU}(\mathbf{x}) \ . \end{aligned} \tag{38}$$

The gradients in Eq. (23) are calculated by applying Eq. (37) and Eq. (38) to Eq. (36).

References

[1] S. J. Price, R. Jena, N. G. Burnet, T. A. Carpenter, J. D. Pickard, and J. H. Gillard, "Predicting patterns of glioma recurrence using diffusion tensor imaging," *Eur. Radiol.*, vol. 17, no. 7, pp. 1675–1684, 2007.

[2] R. Verma, E. I. Zacharaki, Y. Ou, H. Cai, S. Chawla, S.-K. Lee, E. R. Melhem, R. Wolf, and C. Davatzikos, "Multiparametric Tissue Characterization of Brain Neoplasms and Their Recurrence Using Pattern Classification of MR Images," *Acad. Radiol.*, vol. 15, no. 8, pp. 966–977, 2008.

[3] W.-D. Heiss, P. Raab, , and H. Lanfermann, "Multimodality assessment of brain tumors and tumor recurrence," *J. Nucl. Med.*, vol. 52, no. 10, pp. 1585–1600, 2011.

[4] J. M. Provenzale, S. Mukundan, and D. P. Barboriak, "Diffusion-weighted and Perfusion MR Imaging for Brain Tumor Characterization and Assessment of Treatment Response," *Radiology*, vol. 239, no. 3, pp. 632–649, 2006.

[5] A. D. Waldman, A. Jackson, S. J. Price, C. A. Clark, T. C. Booth, D. P. Auer, P. S. Tofts, D. J. Collins, M. O. Leach, and J. H. Rees, "Quantitative imaging biomarkers in neuro-oncology," *Nat. Rev. Clin. Oncol.*, vol. 6, no. 8, pp. 445–54, 2009.

[6] B. L. Dean, B. P. Drayer, C. R. Bird, R. A. Flom, J. A. Hodak, S. W. Coons, and R. G. Carey, "Gliomas: Classification with MR Imaging," *Radiology*, vol. 174, no. 2, pp. 411–415, 1990.

[7] S. Periaswamy and H. Farid, "Medical image registration with partial data," *Med. Image Anal.*, vol. 10, no. 3, pp. 452–464, 2006.

[8] N. Chitphakdithai and J. S. Duncan, "Non-rigid Registration with Missing Correspondences in Preoperative and Postresection Brain Images," in *Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, vol. 6361, 2010, pp. 367–374.

[9] M. Niethammer, G. L. Hart, D. F. Pace, P. M. Vespa, A. Irimia, J. D. V. Horn, and S. R. Aylward, "Geometric Metamorphosis," in *Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, vol. 6892, 2011, pp. 639–646.

[10] O. Clatz, H. Delingette, I.-F. Talos, A. Golby, R. Kikinis, F. A. Jolesz, N. Ayache, and S. K. Warfield, "Robust Nonrigid Registration to Capture Brain Shift From Intraoperative MRI," *IEEE Trans. Med. Imaging*, vol. 24, no. 11, pp. 1417–1427, 2005.

[11] P. Risholm, E. Samset, I.-F. Talos, , and W. Wells, "A Non-rigid Registration Framework That Accommodates Resection and Retraction," in *Inf. Process. Med. Imaging (IPMI)*, vol. 5636, 2009, pp. 447–458.

[12] A. Mohamed, E. I. Zacharaki, D. Shen, and C. Davatzikos, "Deformable registration of brain tumor images via a statistical model of tumor-induced deformation," *Med. Image Anal.*, vol. 10, no. 5, pp. 752–763, 2006.

[13] E. I. Zacharaki, D. Shen, S.-K. Lee, and C. Davatzikos, "ORBIT: A Multiresolution Framework for Deformable Registration of Brain Tumor Images," *IEEE Trans. Med. Imaging*, vol. 27, no. 8, pp. 1003–1017, 2008.

[14] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig, "A brain tumor segmentation framework based on outlier detection," *Med. Image Anal.*, vol. 8, no. 3, pp. 275–283, 2004.

[15] B. H. Menze, K. V. Leemput, A. Honkela, E. Konukoglu, M.-A. Weber, N. Ayache, and P. Golland, "A Generative Approach for Image-Based Modeling of Tumor Growth," in *Inf. Process. Med. Imaging (IPMI)*, vol. 6801, 2011, pp. 735–747.

[16] A. Gooya, K. M. Pohl, M. Billelo, L. Cirillo, G. Biros, E. R. Melhem, and C. Davatzikos, "GLISTR: Glioma Image Segmentation and Registration," *IEEE Trans. Med. Imaging*, vol. 31, no. 10, pp. 1941–1954, 2012.

[17] H. Lamecker and X. Pennec, "Atlas to Image-with-Tumor Registration based on Demons and Deformation Inpainting," in *MICCAI Workshop Comput. Imaging Biomark. Tumors (CIBT)*, 2010.

[18] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios, "Dense image registration through MRFs and efficient linear programming," *Med. Image Anal.*, vol. 12, no. 6, pp. 731–741, 2008.

[19] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Med. Image Anal.*, vol. 12, no. 1, pp. 26–41, 2008.

[20] Y. Ou, A. Sotiras, N. Paragios, and C. Davatzikos, "DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting," *Med. Image Anal.*, vol. 15, no. 4, pp. 622–639, 2011.

[21] G. E. Christensen and H. J. Johnson, "Consistent Image Registration," *IEEE Trans. Med. Imaging*, vol. 20, no. 7, pp. 568–582, 2001.

[22] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, no. 1, pp. S151–S160, 2004.

[23] H. D. Tagare, D. Groisser, and O. M. Skrinjar, "Symmetric non-rigid registration: A geometric theory and some numerical techniques," *J. Math. Imaging Vis.*, vol. 34, no. 1, pp. 61–88, 2009.

[24] A. Sotiras and N. Paragios, "Discrete Symmetric Image Registration," in *IEEE Int. Symp. Biomed. Imaging (ISBI)*, 2012, pp. 342–345.

[25] K. P. Vives and J. M. Piepmeier, "Complications and expected outcome of glioma surgery," *J. Neurooncol.*, vol. 42, no. 3, p. 289302, 1999.

[26] F. W. Kreth, A. Berlis, V. Spiropoulou, M. Faist, R. Scheremet, R. Rossner, B. Volk, and C. B. Ostertag, "The Role of Tumor Resection in the Treatment of Glioblastoma Multiforme in Adults," *Cancer*, vol. 86, no. 10, pp. 2117–2123, 1999.

[27] K. M. Pohl, J. Fisher, S. Bouix, M. Shenton, R. W. McCarley, W. E. L. Grimson, R. Kikinis, and W. M. Wells, "Using the logarithm of odds to define a vector space on probabilistic atlases," *Med. Image Anal.*, vol. 11, no. 5, pp. 465–477, 2007.

[28] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. R. Stat. Soc. Series B.*, vol. 39, no. 1, pp. 1–38, 1977.

[29] C. Hogea, C. Davatzikos, and G. Biros, "An image-driven parameter estimation problem for a reaction-diffusion glioma growth model with mass effects," *J. Math. Biol.*, vol. 56, no. 6, pp. 793–825, 2008.

[30] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes, "A Comparison of Similarity Measures for Use in 2-D-3-D Medical Image Registration," *IEEE Trans. Med. Imaging*, vol. 17, no. 4, pp. 586–595, 1998.

[31] B. Jian and B. C. Vemuri, "A Robust Algorithm for Point Set Registration Using Mixture of Gaussians," in *IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 2, 2005, pp. 1246–1251.

[32] A. S. Roy, A. Gopinath, and A. Rangarajan, "Deformable Density Matching for 3D Non-rigid Registration of Shapes," in *Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, 2007, vol. 4791, pp. 942–949.

[33] J. Lin, "Divergence Measures Based on the Shannon entropy," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 145–151, 1991.

[34] F. Wang, B. Vemuri, and T. Syeda-Mahmood, "Generalized L2-divergence and its Application to Shape Alignment," in *Inf. Process. Med. Imaging (IPMI)*, vol. 5636, 2009, pp. 227–238.

[35] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-posed problems*. W.H. Winston & Sons, 1977.

[36] A. L. Yuille and N. M. Grzywacz, "A mathematical analysis of the motion coherence theory," *Int. J. Comput. Vis.*, vol. 3, no. 2, pp. 155–175, 1989.

[37] M. Nielsen, L. Florack, and R. Deriche, "Regularization, Scale-Space, and Edge Detection Filters," *J. Math. Imaging Vis.*, vol. 7, pp. 291–307, 1997.

[38] V. Lempitsky, S. Roth, and C. Rother, "FusionFlow: Discrete-Continuous Optimization for Optical Flow Estimation," in *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2008, pp. 1–8.

[39] L. Xu, J. Jia, and Y. Matsushita, "Motion Detail Preserving Optical Flow Estimation," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 34, no. 9, pp. 1744–1757, 2012.

[40] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images," *IEEE Trans. Med. Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[41] M. J. Wainwright, T. Jaakkola, and A. S. Willsky, "MAP Estimation Via Agreement on Trees: Message-Passing and Linear Programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3697–3717, 2005.

[42] V. Kolmogorov, "Convergent Tree-Reweighted Message Passing for Energy Minimization," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 28, no. 10, pp. 1568–1583, 2006.

[43] T. W. Sederberg and S. R. Parry, "Free-form deformation of solid geometric models," *ACM SIGGRAPH Comput. Graph.*, vol. 20, no. 4, pp. 151–160, 1986.

[44] T. M. Lehmann, C. Gönner, and K. Spitzer, "Survey: Interpolation Methods in Medical Image Processing," *IEEE Trans. Med. Imaging*, vol. 18, no. 11, pp. 1049–1075, 1999.

[45] M. Bro-Nielsen and C. Gramkow, "Fast Fluid Registration of Medical Images," in *Vis. Biomed. Comput.*, 1996, pp. 267–276.

[46] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Trans. Med. Imaging*, vol. 17, no. 1, pp. 87–97, 1998.

[47] M. Jenkinson, P. R. Bannister, J. M. Brady, and S. M. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *NeuroImage*, vol. 17, no. 2, pp. 825–841, 2002.

[48] S. Smith, "Fast robust automated brain extraction," *Hum. Brain Mapp.*, vol. 17, no. 3, pp. 143–155, 2002.

[49] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics Bull.*, vol. 1, no. 6, pp. 80–83, 1945.

[50] L. R. Dice, "Measure of the amount of ecological association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[51] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. C. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. M. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *NeuroImage*, vol. 46, no. 3, pp. 786–802, 2009.

[52] K. Murphy, B. van Ginneken, J. M. Reinhardt, S. Kabus, K. Ding, X. Deng, K. Cao, K. Du, G. E. Christensen, V. Garcia, T. Vercauteren, N. Ayache, O. Commowick, G. Malandain, B. Glocker, N. Paragios, N. Navab, V. Gorbunova, J. Sporring, M. de Bruijne, X. Han, M. P. Heinrich, J. A. Schnabel, M. Jenkinson, C. Lorenz, M. Modat, J. McClelland, S. Ourselin, S. E. A. Muenzing, M. A. Viergever, D. D. Nigris, D. L. Collins, T. Arbel, M. Peroni, R. Li, G. C. Sharp, A. Schmidt-Richberg, J. Ehrhardt, R. Werner, D. Smeets, D. Loeckx, G. Song, N. J. Tustison, B. B. Avants, J. C. Gee, M. Staring, S. Klein, B. C. Stoel, M. Urschler, M. Werlberger, J. Vandemeulebroucke, S. Rit, D. Sarrut, and J. P. W. Pluim, "Evaluation of Registration Methods on Thoracic CT: The EMPIRE10 Challenge," *IEEE Trans. Med. Imaging*, vol. 30, no. 11, pp. 1901–1920, 2011.

[53] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, "A reproducible evaluation of ANTs similarity metric performance in brain image registration," *NeuroImage*, vol. 54, no. 3, pp. 2033–2044, 2011.

[54] J. Ashburner and K. J. Friston, "Unified segmentation," *NeuroImage*, vol. 26, no. 3, pp. 839–851, 2005.

[55] K. M. Pohl, J. Fisher, J. J. Levitt, M. E. Shenton, R. Kikinis, W. E. L. Grimson, and W. M. Wells, "A Unifying Approach to Registration, Segmentation, and Intensity Correction," in *Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, vol. 3749, 2005, pp. 310–318.

[56] K. M. Pohl, J. Fisher, W. E. L. Grimson, R. Kikinis, and W. M. Wells, "A Bayesian model for joint segmentation and registration," *NeuroImage*, vol. 31, no. 1, pp. 228–239, 2006.

[57] K. V. Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated Model-Based Bias Field Correction of MR Images of the Brain," *IEEE Trans. Med. Imaging*, vol. 18, no. 10, pp. 885–896, 1999.

[58] D. Kwon, K. J. Lee, I. D. Yun, and S. U. Lee, "Nonrigid Image Registration Using *Dynamic* Higher-Order MRF Model," in *Eur. Conf. Comput. Vis. (ECCV)*, vol. 5302, 2008, pp. 373–386.

[59] D. Kwon, I. D. Yun, K. M. Pohl, C. Davatzikos, and S. U. Lee, "Nonrigid Volume Registration Using Second-Order MRF Model," in *IEEE Int. Symp. Biomed. Imaging (ISBI)*, 2012, pp. 708–711.

[60] D. Rueckert, P. Aljabar, R. A. Heckemann, J. V. Hajnal, and A. Hammers, "Diffeomorphic Registration Using B-Splines," in *Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, 2006, pp. 702–709.

[61] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A Comparative Study of Energy Minimization Methods for Markov Random Fields," in *Eur. Conf. Comput. Vis. (ECCV)*, vol. 3952, 2006, pp. 16–29.