Published in final edited form as:

IEEE Trans Med Imaging. 2021 January ; 40(1): 193–204. doi:10.1109/TMI.2020.3024264.

Fully Automatic Calibration of Tumor-Growth Models Using a Single mpMRI Scan

Klaudius Scheufele, Shashank Subramanian, George Biros [Member, IEEE]

Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712 USA

Abstract

Our objective is the calibration of mathematical tumor growth models from a single multiparametric scan. The target problem is the analysis of preoperative Glioblastoma (GBM) scans. To this end, we present a fully automatic tumor-growth calibration methodology that integrates a single-species reaction-diffusion partial differential equation (PDE) model for tumor progression with multiparametric Magnetic Resonance Imaging (mpMRI) scans to robustly extract patient specific biomarkers i.e., estimates for (i) the tumor *cell proliferation rate*, (ii) the tumor *cell migration rate*, and (iii) the original, *localized site(s) of tumor initiation*. Our method is based on a sparse reconstruction algorithm for the tumor initial location (TIL). This problem is particularly challenging due to nonlinearity, ill-posedness, and ill conditioning. We propose a coarse-to-fine multi-resolution continuation scheme with parameter decomposition to stabilize the inversion. We demonstrate robustness and practicality of our method by applying the proposed method to clinical data of 206 GBM patients. We analyze the extracted biomarkers and relate tumor origin with patient overall survival by mapping the former into a common atlas space. We present preliminary results that suggest improved accuracy for prediction of patient overall survival when a set of imaging features is augmented with estimated biophysical parameters. All extracted features, tumor initial positions, and biophysical growth parameters are made publicly available for further analysis. To our knowledge, this is the first fully automatic scheme that can handle multifocal tumors and can localize the TIL to a few millimeters.

Keywords

Glioblastoma; model calibration; tumor-growth inversion; tumor initiation

I. Introduction

Glioblastoma multiforme (GBM) is a highly aggressive primary brain tumor with poor prognosis despite radical treatment. GBMs are challenging due to their heterogeneity and their infiltration into surrounding healthy tissue. Several research groups are working on developing image-driven biophysical tumor-growth models for GBMs that can be applied in a clinical setting and assist with diagnosis, prognosis, and treatment planning. For example,

(Corresponding author: Klaudius Scheufele.) klaudius.scheufele@gmail.com.

This article has supplementary downloadable material available at <https://ieeexplore.ieee.org>, provided by the authors.

such models can be used to determine survival, probability of recurrence, molecular subtypes, overall tumor aggressiveness, and epidemiology [1]–[4].

In this paper, we focus on the problem of calibrating a tumor-growth model from a single scan. Indeed, for the majority of GBM cases, preresection clinical decisions are based only on a single scan. Our reconstruction produces biomarkers that can potentially offer additional diagnostic and prognostic value, as they might relate to the tumor's aggressiveness and extent of infiltration beyond imaging-visible tumor boundaries. Given a multiparametric, tumor-bearing, pretreatment scan, we want to calibrate various parameters in a biophysical model. (This is the so-called *inverse problem*.) We are particularly interested in the *tumor initiation location (TIL)*. Determining the TIL is mathematically required for robustly identifying tumor growth parameters (cf. Fig. 2). Moreover, the tumor location could be related to distinct tumor mutations, so it could add complementary information to radiogenomics based analyses [6]–[8]. The observed tumor location plays a key role in guiding surgical and radiation treatment, and has prognostic value for patient survival [6], [9]. *Our hypothesis is that TIL could play a similar role. Here, we are not performing an exhaustive test for this hypothesis. Instead, we focus on introducing an algorithm to recover the TIL and on performing a preliminary analysis of its correlation with survival.* Although several methods have been proposed for estimating tumor parameters, there is surprisingly little work on robustly estimating the TIL.

A. Contributions

After an image segmentation preprocessing step, we solve an inverse problem for two scalar tumor parameters (*migration rate* and *proliferation rate*), and the *TIL*, which is parameterized by a high-dimensional vector. In our recent work [10], we presented theoretical results on the non-convexity and ill-posedness of the problem, introduced an overall regularization framework, and verified the method on synthetic datasets. However, our method in [10] is not robust enough to be applied to generic clinical datasets. Here, we expand [10] in several ways: • We stabilize and accelerate the solution of the inverse problem by introducing a (*multilevel*) *coarse-to-fine* continuation scheme, and embed a two-step optimization scheme based on parameter decomposition. This is critical to prevent the method from producing infeasible or sub-optimal solutions. • We enhance [10] to a *fully automatic*, and *robust* method for GBM biomarker extraction from actual clinical imaging data. We demonstrate robustness across 206 patients from the BraTS dataset [11], featuring vastly heterogeneous GBM phenotypes along with multifocality and multicentricity. The extracted features are available in the supplementary material (SM). We correlate our results with a purely imaging based method to identify multifocal/multicentric tumor seeds (using connected components and center of mass). • We create a *GBM atlas of TIL* from the 206 case-cohort via atlas-registration (see Fig. 1). This helps to identify high-risk tumor sites and correlations between the TIL and the severity and progression of the disease. We present preliminary results for patient survival predictions based on extracted GBM biomarkers.

B. Related Work

Mathematical modeling in cancer progression has a long tradition; in particular forward models for tumor progression are strongly represented in literature [3], [12], [13]. There is

less work on patient-specific model calibration. Similar to our approach, most groups rely on reaction-diffusion models and calibrate parameters using MRI scans [12], [14]–[17]. Despite its simplicity, this framework has been shown to be effective in medical MR image analysis, segmentation, and tumor characterization [3], [12], [18]–[20].

Several parameter calibration studies assume access to subject-specific longitudinal imaging data (from two or more time points), which are rarely available for GBMs. In contrast, our approach requires one single pretreatment scan only. Besides our work in [10], approaches for single-snapshot data have been presented in [3], [19], [21] (and the references therein for older works). The groups in [19], [21] integrate image registration with tumor inversion methods; Scheufele *et al.* [21] use an L_2 penalty that results in non-localized TIL (see Fig. 1) that lead to erroneous predictions of infiltration and proliferation; Gooya *et al.* [19] manually select tumor seeds, and refine the location TIL using EM. In our approach, the number of seeds is not preselected but it is automatically determined by our algorithm, which is particularly important for multifocal tumors. The works in Menze *et al.* [22] and Lipkova *et al.* [3] employ a Bayesian framework for parameter calibration, but restrict tumor initiation to only one distinct location, which limits practical applicability to multifocal pathologies. For more discussion, see [10].

Tumor location carries important information for low and high-grade gliomas [8], [23]. For GBMs, Le *et al.* [9] analyzed the correlation between tumor location and clinical properties of glioblastomas in frontal and temporal lobes for 406 patients; Ren *et al.* [7] analyzed 528 patients with gliomas with scans and histological data and analyzed incidences of 1p/19q co-deletion and IDH1/2 mutation in gliomas by regions and hemispheres in the brain. Roux *et al.* [6] found longer survival in those patients with tumors in the left temporal lobe regardless of whether the patient received resection or biopsy; moreover, they found distinct spatial distributions of tumor determined by molecular biomarkers like MGMT. Binder *et al.* [24] correlated the shape of tumor with EGFR mutations, and Bilello *et al.* [8] correlated such mutations with the observed tumor location. Further evidence suggesting that the location of tumor initiation relates to distinct tumor mutations and might correlate with patient survival can be found in [5], [20], [25]–[28].

C. Outline

In §II, we introduce the model and the new calibration algorithm. In §III, we present results from the BraTS dataset and analyze correlations to patient survival. More technical details and results can be found in the supplementary material.

II. Methods

A. Tumor Growth Model

1) Model and Materials: Our tumor-growth model is a standard single-species nonlinear reaction-diffusion partial differential equation (PDE) [29]

$$\partial_t c - \kappa \mathcal{D}c - \rho \mathcal{R}c = 0 \quad \text{in } \Omega_{\mathcal{B}} \times (0, 1], \quad (1a)$$

$$c(0) = c_0 \text{ in } \Omega_{\mathcal{B}}, \quad (1b)$$

Here $c = c(\mathbf{x}, t) \in [0, 1]$ denotes the tumor concentration evolving over time $t \in [0, 1]$ in the three-dimensional brain domain $\Omega_{\mathcal{B}} \subset \Omega$, embedded into the simulation domain, $\Omega = [0, 2\pi]^3$. Tumor invasion and proliferation are modeled by the *diffusion* operator $\mathcal{D}c = \text{div } \kappa_m \nabla c$, and a logistic *reaction* term $\mathcal{R}c = \rho_m c(1 - c)$, respectively, with spatially variable coefficients ρ_m and κ_m defined based on the underlying material properties $\mathbf{m}$. The latter indicate a vector of voxelwise probabilities m_i for *white matter* (W), *gray matter* (G), and *cerebrospinal fluid with ventricles* (F), which are inferred from a deep convolutional neural network [30] using mpMRI. Tumor growth is assumed along the white matter fiber tracts only. We denote characteristic cell proliferation and cell migration in white matter by ρ_w and κ_w , respectively. Tumor cells do not grow or invade into the cerebrospinal fluid-filled ventricles. Details and examples of tumor growth for different values for ρ_w and κ_w are given in the supplementary materials. Although the time horizon is not known, in [10] we show that, after non-dimensionalization, only κ and ρ are independent variables and we can set the time horizon to one by simple rescaling. We approximate noflux boundary conditions $\partial\Omega_{\mathcal{B}}$ via a penalty approach [31], and use periodic boundary conditions on Ω . Finally, $c(0) = c(\mathbf{x}, 0)$ is the initial concentration of the tumor c_0 . Our hypothesis is that it is highly localized in space.

Following [10], we introduce a parametrization of c_0 as follows

$$c_0 = \Phi^T(\mathbf{x})\mathbf{p} = \sum_{i=1}^{n_p} \Phi_i(\mathbf{x})p_i, \quad \mathbf{p} \in \mathbf{R}^{n_p}, \quad (2)$$

where $\Phi(\mathbf{x}) = \{\Phi_i(\mathbf{x})\}_{i=1}^{n_p}$ is a set of Gaussian basis functions with $\Phi_i(\mathbf{x}) = \Phi(\mathbf{x} - \mathbf{x}_i, \sigma)$, centered at location $\mathbf{x}_i$. The standard deviation σ controls the width of the Gaussian basis functions. Here, we make the assumption that the spatial distribution of the initial tumor is given by Φ_i for some i , and sufficiently small σ .¹ We invert for

$$\mathbf{g} = (\rho_w, \kappa_w, \mathbf{p}), \quad (3)$$

i.e., the tumor *proliferation* (a scalar) and *infiltration* (a scalar) rate in white matter, and the *TIL* (a vector) parameterized by $\mathbf{p}$.

2) Calibration and Extraction of Biomarkers: We calibrate biophysical parameters $\mathbf{g}$ so that the predicted tumor evolution of the model in (1) matches the patient segmentation. As input, the inverse solve takes the mpMRI segmentation [30] and derives probability/cell concentration maps m_i , $i = W, G, F$ for the healthy tissue anatomy labels and for the tumor sub-regions, i.e., m_T for the tumor core (necrotic core, enhancing tumor, and non-enhancing

¹The continuum hypothesis of our model breaks down at the microscopic scale of tumor initiation. Thus, we use our PDE model after the tumor has grown sufficiently large

tumor), and m_E for the peritumoral edema; detailed in §II-C.1. Then, we solve the PDE-constrained nonlinear least squares problem

$$\min_{\mathbf{g}} \mathcal{J}(\mathbf{g}) = \frac{1}{2} \|\mathcal{O}(c(1) - d)\|_{L_2(\Omega)}^2 + \beta \mathcal{S}(\mathbf{g}) \quad (4a)$$

$$\text{subject to } \mathcal{F}(\mathbf{g}) \text{ given by (1), and } \mathcal{E}(\mathbf{g}), \quad (4b)$$

minimizing the discrepancy between observed model prediction $c(1)$ and target data $d = m_T$ in a least squares sense, balanced with a regularization term $\mathcal{S}(\mathbf{g})$ (with weight β) to alleviate ill-conditioning. $\mathcal{E}(\mathbf{g})$ is a set of additional constraints, allowing us to *simultaneously invert* for the set of parameters $\mathbf{g}$, and promote *sparse localization* of the TIL. Data discrepancy is only measured at certain *observation points*, defined by an **observation operator** $\mathcal{O}$, accounting for the fact that the true extent of tumor infiltration is not directly observable. This operator defines a region with clearly observable tumor margin (while the true tumor may extend beyond this margin); details are given in §II-B.2. We introduce a robust algorithm for the solution of the inverse problem by employing a multilevel continuation scheme based on spatial grid-coarsening. This reduces the risk of early termination in (bad) local minima, and improves the practicality of the method for real clinical data, featuring single- and multifocal pathologies with great heterogeneity. Although we used and L_2 similarity in (4), our framework can be used with other similarity functions, for example cross-entropy.

Reconstructing these parameters is a hard problem. It is well known that inversion for all parameters of a reaction-diffusion problem is ill-posed, since the forward model $\mathcal{F}$ is non-unique in the sense that different parameter configurations produce the same solution [10], [16]. Thus, additional constraints need to be imposed to alleviate the ill-posedness. We enforce $\max_{\mathbf{x}} c_0(\mathbf{x}) = 1$, corresponding to the modeling assumption that at $t = 0$ the tumor volume concentration reaches 100% of a voxel for the first time. Additionally, we need to ensure the tumor initial condition to be small and localized. We impose an ℓ_0 constraint $\|\mathbf{p}\|_{\ell_0(\mathbf{R}^n p)} \leq \nu$, restricting the initial tumor to consist of a small number ν of initiations (i.e., active Gaussians) only. For numerical optimization, we introduce additional bound constraints on the inversion variables to safeguard against infeasible solutions and bad local minima. We complete the formulation by defining the regularization term $\mathcal{S}(\mathbf{g})$ as $\frac{1}{2} \|\Phi^T \mathbf{p}\|_{L_2(\Omega)}^2$, a standard Tikhonov regularization model.

Hence, our **inverse problem** reads:

$$\min_{(\rho_w, \kappa_w, p)} \frac{1}{2} \|\mathcal{O}(c(1) - d)\|_{L_2(\Omega)}^2 + \frac{\beta}{2} \|\Phi^T \mathbf{p}\|_{L_2(\Omega)}^2 \quad (5a)$$

$$\text{s. t. } \begin{cases} \mathcal{F}(\rho_w, \kappa_w, \mathbf{p}) & \text{given by (1),} \\ \max \Phi^T \mathbf{p} = 1 & \text{in } \Omega, \\ \|\mathbf{p}\|_{\ell_0(\mathbf{R}^{n_p})} \leq \nu & \text{in } \mathbf{R}^{n_p}, \\ 0 \leq \mathbf{p} & \text{in } \mathbf{R}^{n_p}, \\ \kappa_{min} \leq \kappa_w \leq \kappa_{max} & \text{in } \mathbf{R}, \\ \rho_{min} \leq \rho_w \leq \rho_{max} & \text{in } \mathbf{R}. \end{cases} \quad (5b) (5c) (5d) (5e) (5f) (5g)$$

Next, we discuss the numerical solution of (5).

B. Numerical Scheme

In §II-B.1, we summarize the scheme from [10] for solving (5). Then, in §II-B.2 we discuss the novel algorithmic modifications to enhance robustness in terms of avoiding stagnation, divergence of the optimizer, arbitrary parameter selection, and challenges with multifocal tumors.

1) Biophysical Parameter Inversion:

a) Selection of Gaussian Basis Functions: The set $\{\Phi_i(x)\}_{i=1}^{n_p}$ of Gaussian basis functions is determined automatically; that is, the number n_p and spatial distribution (i.e., centers $\{x_i\}_{i=1}^{n_p}$) are defined based on an input tumor concentration map m_Φ , e.g., tumor core input data $m_\Phi = m_T$, or the tumor initial condition $m_\Phi = c(0)$. We refer to the supplementary materials for details.

b) Enforcing the Sparsity Constraint: Solving the L_2 -regularized problem (5a)–(5b) typically results in tumor initial conditions which closely resemble the input data $d = m_T$ up to scaling. As a result, the parameters ρ_w and κ_w are underestimated [10]. The sparsity constraint (5d) restricts TIL to a few, small, and localized positions, i.e., the solution $\mathbf{p}$ is constrained to have at most ν non-zero entries. The corresponding active Gaussians form the support of $\mathbf{p}$ or $\text{supp}(\mathbf{p})$, denoting the set of non-zero indices of $\mathbf{p}$. We employ an iterative recovery strategy based on compressive sampling [32], and determine important components in $\mathbf{p}$ based on the gradient of the objective (5a) (i.e., large gradient corresponds to large data-misfit). A solution to (5) is found iteratively by alternating between (i) identifying a set of components $\mathcal{V}$, $|\mathcal{V}| \leq 3\nu$ using the gradient, and (ii) solving an L_2 -regularized problem (5a)–(5b), (5e)–(5f) for $\mathbf{g}$ on these components. That is, we iterate:

1. *Identify Significant Gaussians*, i.e., Gaussians corresponding to the 2ν largest components of $\nabla \mathcal{J}$.
2. *Merge Support* of current solution with new Gaussians from step 1; this forms $\mathcal{V}$.
3. *Find Solution $\mathbf{g}$* of (5a)–(5b), (5e)–(5f) within the subspace defined by $\mathcal{V}$.
4. *Restrict Solution* to ν largest components in $\mathbf{p}$, and set $p_i = 0$ for remaining components.

The details can be found in [10].

c) Enforcing the Max Constraint: We implement the max constraint $\max \Phi^T \mathbf{p} = 1$ in Ω by following a two-step strategy, keeping either ρ_w or $\mathbf{p}$ fixed:

- I. ℓ_0 -phase: Given an estimate for ρ_w , we solve steps 1–4 from above for parameters $(\kappa_w, \mathbf{p}) \subset \mathbf{g}$, and keep ρ_w constant during this phase.
- II. ρ/κ -phase: Given an estimate for $(\kappa_w, \mathbf{p})$, we rescale $\mathbf{p}$ such that $\max \Phi^T \mathbf{p} = 1$ is fulfilled, and solve (5) for parameters $(\kappa_w, \rho_w) \subset \mathbf{g}$, keeping the (rescaled) vector $\mathbf{p}$ fixed.

d) Numerical Optimization of the L_2 -Regularized Problem: We solve (5a)–(5b), (5e)–(5f) in an optimize-then-discretize fashion using an adjoint-based approach to compute gradients. The resulting optimization problem is solved using a bound constrained, limited memory BFGS quasi-Newton solver globalized with Armijo line-search. Details on the numerical solution of the forward problem (1) are given in the supplementary material.

Next, we describe how (5) is solved using a *coarse-to-fine multigrid approach with parameter decomposition*, to improve robustness and practicality of the scheme.

2) A Robust Continuation Scheme With Parameter Decomposition: We stabilize and speed up the solution of (5) by a *multilevel coarse-to-fine continuation scheme*. We consider image resolutions $n^\ell \in \{64^3, 128^3, 256^3\}$, $\ell = 1, 2, 3$, and solve a series of three optimization problems, where the coarse-level solution serves as an initial guess, increasing the chance of keeping the Newton method within the basin of attraction of the global minimum. Besides coarsening the spatial resolution, we also coarsen the bandwidth $\sigma^\ell \rightarrow 2\sigma^\ell = \sigma^{\ell-1}$ of the Gaussian basis functions across levels. On levels 128^3 and 256^3 , we select the Gaussians basis functions based on the coarse level solution for the tumor initial condition, i.e., we set $m_\Phi := \alpha(0)^{\ell-1}$, while on the coarsest level 64^3 , we use the tumor core input data $m_\Phi := m_T$ for Gaussian selection. This strategy allows to robustly refine to a very small target bandwidth of $\sigma = 2\pi/256$ on the original image resolution $n^3 = 256^3$, to enable very localized (point-like) tumor initial conditions. Furthermore, high resolution is required to reliably invert for the diffusion coefficient κ_w [10].

a) Injection of Coarse Level Solution $\mathbf{g}^{\ell-1}$: Parameters $\alpha(0)$ and ρ_w are correlated: a non-localized initial condition results in underestimation of ρ_w (cf. supplementary material). Changing the bandwidth σ across levels changes the scale of ρ_w . To compensate this, we modify the max constraint (5c) to $\max \Phi^T \mathbf{p} = \delta^\ell$ with $\delta^\ell \in \{1/4, 1/2, 1\}$ for $\ell = 1, 2, 3$, such that the integral $\int_\Omega \Phi \, dx = \text{constant}$, across levels. The estimation of $\rho_w^{\ell-1}$ and $\kappa_w^{\ell-1}$ on the coarse level is then used as initial guess on the fine level. Similarly, $\mathbf{p}^{\ell-1}$ is injected as initial guess on the fine level.

b) Enabling Multifocality/Multicentricity: The “sparsity level” ν of the initial tumor $\alpha(0)$ is an important parameter. If ν is too large the TIL is not localized, which is not physical and leads to wrong growth parameters. If ν is too small we cannot reconstruct complex tumor shapes, for example multifocal tumors. To facilitate an *a priori* choice for

ν , we perform a connected component analysis (see supplementary materials for details) on the tumor input data m_T , and define a target sparsity $\nu_{c,j}$ per component $j \in \{1, \dots, n_c\}$ as follows:² We set the overall sparsity of $\alpha(0)$ to $\nu = 5n_c$, and sub-divide it into a *per-component sparsity* by attributing 3 DOFs to each component, and distributing the remaining $2n_c$ DOFs according to the relative mass.³ Steps 1–4 in the ℓ_0 -phase are then performed as per-component. This strategy allows for robustly dealing with multifocal tumors of vastly heterogeneous size and shape.

The selection of the parameters was verified using synthetic data and numerical experiments. We discuss the details in the supplementary material. The overall calibration however can be considered parameter-free in the sense that all employed parameters are *fixed* across runs.

c) Observation operator: The *observation operator* $\mathcal{O}$ appears in the L_2 mismatch term in (5a) and controls which voxels we select to compare prediction and data and the weight of the comparison. This observation operator is a modeling decision but appears only in the calibration problem and has a significant effect in solution of (5a) (see Fig. 3). In particular it is given by

$$\mathcal{O}(x) := \mathbb{1}_{[x \in TC]} + \lambda \mathbb{1}_{[x \in B \setminus WT]}, \quad \lambda \in [0, 1], \quad (6)$$

with the indicator function $\mathbb{1}_{[\cdot]}$. Here, $B \setminus WT$ indicates the complement of the area occupied by the whole tumor (WT; defined as TC+ED). Let us now explain this formula.

The need for $\mathcal{O}$ is due to the fact that the true extent of tumor infiltration cannot be observed from MRI. Moreover, probability maps of tissue labels inferred from DNN for MRI segmentation are typically rather sharp and almost binary. A common way to integrate such data with inverse tumor modeling is to take the segmentation of tumor core as probability (or tumor cell density) of one and fit the model prediction in an L_2 sense. This is, however, not a realistic representation of the heterogeneous tumor structure and sub-regions. In particular the peritumoral edema (an inflammatory region surrounding the tumor), typically presents a higher concentration of malignant cells, yet cannot be considered cancer tissue. Our tumor growth model is lacking a description for the formation of edema. For the above approach (where the data is matched everywhere), to capture the formation of edema, we would have to define a value for the tumor concentration within this region, which we want the model to match. Such a value is not known.

Instead of using all the voxels to define the data mismatch

$$\frac{1}{2} \sum_{x \in \mathcal{B}} (c(1) - d)^2,$$

² n_c is the number of components with relative mass above 0.1%, as obtained from connected component analysis. This threshold is to discard very small “components” caused by segmentation errors (typically one or few connected voxels); their relative size (to the brain volume) is in the order of 0.001% and smaller, whereas for significant components relative sizes are between 100% and 5%. Our results are not sensitive to the 0.1% value. See supplementary materials for details.

³To account for potentially erroneous segmentations used to derive the target data $d = m_T$, we additionally attribute one degree of freedom to components with relative mass below 0.1%.

we only use a subset of voxels (partial observaions), by defining the data mismatch as

$$\frac{1}{2} \sum_{x \in TC} (c(1) - d)^2 + \frac{\lambda}{2} \sum_{x \in \mathcal{B} \setminus WT} (c(1) - d)^2. \quad (7)$$

The data mismatch term using $\mathcal{O}$ is identical to (7). In other words, this operator results in an inverse model where tumor cell densities (i) are matched to the data within TC, (ii) are not matched to the data within ED, and (iii) are matched, but discounted by λ within the region beyond the visible edema. As a result, predicted tumor cell density within the region of peritumoral edema is naturally governed by the model dynamics, and not directly fitted to an arbitrary value. This is different from most common approaches. Finally, the parameter λ controls how much predicted tumor cell concentraton outside the visible tumor with edema should be penalized (since the target data is zero there). It can be interpreted as uncertainty of observation in this region. A value of $\lambda = 1$, e.g., means no uncertainty in the target data, and predicted tumor concentration outside WT is penalized with full weight; on the other extreme, a value of $\lambda = 0$ would discount errors in this region to zero (a model prediction of one in the entire domain, would, e.g., yield zero mismatch for this choice). We experimentally assess the performance for different values of λ in §III-A. Fig. 3 compares the reconstruction of the tumor cell density from DNN segmentations for different choices of the observation operator. In particular, a one-dimensional profile of TC and ED segmentation, overlaid with the reconstructed tumor cell desities illustrates the effect of different choices of lambda as opposed to observing the data as is (everywhere).

Algorithm 1

Outline of Inversion Scheme Using Grid-Continuation and Parameter Decomposition in Pseudocode. During the ρ/κ -**phase**, the Vector $\mathbf{p}$ is Fixed, and Estimates for ρ_w and κ_w Are Computed. In the ℓ -**Phase**, ρ_w is Held Constant to solve for a Sparse $\mathbf{p}$, and κ_w , **concomp** performs a connected component analysis on m_T (See Supplementary Materials for Details), Returning the Number of Components n_c , and the Sparsity Per Component $\mathbf{v}_c$.

Input : Initial guess $\mathbf{g}^0$, and label maps $\mathbf{m}$, m_T , m_E

Output : Solution $\mathbf{g} = (\rho_w, \kappa_w, \mathbf{p})$, $\alpha(0)$

```

1  down-sample  $\{\mathbf{m}\}_{\ell}^3 = 1, \{m_T\}_{\ell}^3 = 1, \{m_E\}_{\ell}^3 = 1$ 
2   $m_{\Phi} \leftarrow m_T^{\ell} = 1$ 
3  for  $\ell = 1, 2, 3$  do
4    select Gaussians based on  $m_{\Phi}$ 
5     $\mathcal{O} := \mathbb{1}_{[x \in TC]} + \lambda \mathbb{1}_{[x \in B \setminus WT]}, \sigma^{\ell} = 2\pi/n^{\ell}, \mathcal{V} \leftarrow \{\}$ 
6     $v_c, n_c, c \leftarrow \text{concomp}(m_T^{\ell})$ 
7    if  $\ell > 1$  then
8       $\mathbf{g}^{\ell} \leftarrow \mathbf{g}^{\ell-1}(\text{injection}), \mathcal{V} \leftarrow \text{supp}(\mathbf{p}^{\ell-1})$ 
9      rescale  $\mathbf{p}$  s.t.  $\max \Phi^T \mathbf{p} = n^{\ell}/256$ 
10      $\rho_w, \kappa_w \leftarrow \rho/\kappa\text{-phase} \# \mathbf{p} = \text{fixed}$ 
```

```

11   for  $k = 1, \dots, \iota$  do  $\ell$ -phase
12      $\mathcal{V} \leftarrow \text{supp}(p) \cup \text{supp}(|\nabla \mathcal{F}(g)|^{2\nu})$ 
13     solve (5) in  $\mathcal{V}$  for  $(\kappa_w, p) \# \rho_w = \text{fixed}$ 
14      $p \leftarrow p|_{\mathcal{V}(\text{restriction})}$ ,  $\mathcal{V} \leftarrow \text{supp}(p)$ 
15     solve (5) in  $\mathcal{V}$  for  $(\kappa_w, p) \# \rho_w = \text{fixed}$ 
16     rescale  $p$  s.t.  $\max_{\Phi} p = n/256$ 
17      $\rho_w, \kappa_w \leftarrow \rho/\kappa$ -phase  $\# p = \text{fixed}$ 
18     up-sample  $c(0)^\ell$  and set  $m_{\Phi} \leftarrow c(0)^\ell$ 

```

d) Summary of Inversion Scheme: We summarize the overall scheme in Alg. 1. The grid-continuation strategy provides good initial guesses for the inversion parameters ρ_w , κ_w , and p . On the coarsest level, Gaussians are selected based on the input data $d = m_T$, while on subsequent levels we use the coarse level initial condition. If a coarse solution is available, we invert for parameters ρ_w and κ_w (line 10), before seeking for a sparse initial condition (lines 11–14). This stabilizes the inversion for the location of tumor initiation, which is largely affected by ρ_w and κ_w . On every level, connected component analysis is performed on the data $d = m_T$ (line 6), defining the sparsity level $\nu = \{v_{c_j}\}_{j=1}^{n_c}$ per component. Gradient and solution thresholding (lines 12 and 14) is then performed *per component*, i.e., the largest v_{c_j} entries from $\nabla \mathcal{F}(g)$ and p are selected, respectively.

e) Remarks on Robustness and Convergence: The discussed extensions to the original inversion scheme [10] are inevitable, to cope with complicated multifocal tumors and great heterogeneity when targeting real clinical data. In fact, if omitted, the inversion diverges or terminates in bad local minima. In particular, a good initial guess for g is critical to keep the optimizer within a basin of attraction of the solution, and helps to cope with the severe non-convexity of the problem. Further, we found that (incorrect) values for ρ_w and κ_w largely affect the inversion for a sparse initial condition, and vice versa and, thus, the parameter decomposition scheme has to be chosen carefully. Together, two concepts deliver the required robustness: (i) the coarse level initial guess, and (ii) the choice of parameter decomposition. For the latter, it is important to solve for ρ_w and κ_w *before* and *after* inversion for p , and to only keep ρ_w and p alternately fixed, but always invert for κ_w .

3) Atlas Registration: To analyze our cohort of 206 patients, we co-registered them to a template using our large-deformation diffeomorphic image registration software CLAIRE [33]. Patient images are first affinely registered to the atlas space, followed by large-deformation, diffeomorphic registration of tissue label maps m_W , m_G , and m_F using CLAIRE. Since we do not have a mpMRI of the healthy brain of the patient, we approximate it using the tumor-bearing scan in which edema and tumor regions in the current scan are replaced by white matter. More sophisticated methods that account for mass-effect exist, but here we opted for a simpler solution. The reconstructed locations of tumor initiation are then mapped and superimposed in a common atlas space.

C. Workflow Summary

Before evaluating our methodology, let us summarize some remaining implementation details.

1) Imaging Data Processing: We consider clinical 3D MRI data of 206 GBM patients from the BraTS'18 training data set [11]. The dataset provides expert segmentation of different GBM phenotypes into necrotic tumor core (NEC), enhancing tumor (EN), and peritumoral edema (ED). In a pre-processing step, we infer binary segmentations of healthy tissue labels, i.e., white matter (WM), gray matter (GM), cerebrospinal fluid (CSF), and ventricles (VE) (or, if not available segmentations of tumor regions, respectively) from mpMRI data, using a deep convolutional neural network [30]. The observations used for calibration are the tumor core (TC), composed of the labels (NEC) and (EN). Solver settings and parameter choices are given in the supplementary material.

III. Results

We examine robustness of our methodology against 206 GBM cases from the BraTS'18 training cohort, featuring great inter-subject variability and vastly heterogeneous GBM phenotypes. We report distributions of calibrated model parameters and other image based features. For 161 cases, we also have patient age, treatment status, and patient overall survival. We investigate correlations between image-based and biophysical model features (such as tumor origin, growth rates, tumor size, and others) with survival. In summary, we report the following:

- *Inversion* (§III-B).
Tab. I: proliferation and migration rates; reconstruction errors; statistics for the BraTS'18 dataset. Fig. 4: reconstruction of the OT and TIL; comparison with center of mass of tumor components; variability of several global parameters across BraTS'18 subjects.
- *Atlas construction* (§III-C).
Fig. 5: localized TILs for all BraTS'18 subjects in atlas space (TIL probability atlas).
- *Survival analysis* (§III-D).
Tab. II: results using hand-crafted features. Fig. 6: regional correlations using TIL and OT probability atlases.

In the supplementary material (SM) we report the following: examples of sensitivity of solution to tumor parameters; sensitivity to the observation operator (§II-B.2); examples of TIL localization; Atlas TILs grouped per survival class; detailed regional analysis survival results; and all reconstruction results for the 206 BraTS'18 subjects.

Before discussing the results in detail, let us introduce some notation. To compare the prediction of the biophysical model ($c(1)$) to the observed tumor segmentation (m_T), we compute $\mu = \|c(1) - m_T\|_{L^2(m_T)}^2 / \|m_T\|_{L^2(m_T)}^2$.⁴ With $\int_X c_1$, we denote integrals over

subregions $X \in \{TC, ED, B \setminus WT\}$. These are used for feature extraction. D_{cm} is the distance between the center of mass (CM) of $c(0)$ (TIL), and the CM of the tumor input data m_T (in voxels);⁵ D_p is the distance between the two most weighted locations x_{p1} and x_{p2} of TIL. D_{cm} and D_p are used to quantitatively assess variability of estimated TIL and, in particular, the deviation from the CM of the input data.

A. Sensitivity to the Observation Operator

We conduct numerical experiments for varying choices of $\lambda \in [0, 1]$ in the observation operator (6), introduced in §II-B.2, and study the sensitivity of the inverse scheme to the latter. In particular, we compare our choice of partial observations to a full observation in the entire brain. We judge quality by the ability to recover plausible tumor cell densities for both, tumor core area as well as for the region of peritumoral edema. At the same time, predicted tumor cell densities for the remaining healthy brain should be small or zero.

1) Results: Quantitative tumor inversion results are given in Tab. 3 for full observation and partial observations (for varying λ), along with qualitative results for two individuals. More detailed results can be found in the supplementary material.

2) Discussion: We observe sensitivity of the estimated growth parameters ρ_w and κ_w to the choice of the observation operator. The scale of sensitivity varies across individuals. While κ_w generally increases for smaller λ , ρ_w seems to be less sensitive; for both parameters, however, we also observe large deviations (see Fig. 3). If no observation operator is considered (panels c) and f) in Fig. 3), the model predicts high tumor cell density values for only a fraction of the TC area in the data (indicated as white dashed line). In particular, predicted cell density does not extend far beyond the TC region, and tumor concentration in the edema region is very low (cf. measure $\int_{ED} c_1$). Using the observation operator in (6), we get a more realistic representation. For smaller values of λ , data-misfit outside the whole tumor area (i.e., tumor prediction in healthy tissue, given by $\int_{B \setminus WT} c_1$) is discounted, resulting in more spread out tumor patterns, reaching far beyond the visible TC region (cf. panels b) and e)). The quality of reconstruction for different choices of λ is hard to quantify (since the extend of tumor cell invasion into preilesional edema is unknown). From our experiments, we find that $\lambda = 1$ shows overall good outcomes and balances the reconstruction of tumor core and edema. Further, simply not observing values for the region of edema is a more natural choice than additionally discounting data mismatch within the healthy tissue. We therefore choose $\lambda = 1$, i.e., $\mathcal{O}(x) := \mathbb{1}_{[x \in B \setminus ED]}$ as observation operator.

B. TIL Computation for BraTS'18 High Grade Gliomas

1) Results: We report representative results in Tab. I. We qualitatively examine (i) the sparse localization and spatial distribution of TIL, as well as (ii) the reconstruction ability and robustness when given complicated shapes, and large, multifocal tumors by visual inspection of various 2D slices given in Fig. 4. We report statistics on the data mismatch

⁴We prefer this over a DICE coefficient, which would require arbitrary thresholding of $c(1)$.

⁵The distance is computed per (tumor) component (see §II-B). We only report the distance for the largest component here. The distance for all components can be found in the supplementary material.

and distance of TILs in Tab. I from 206 cases. Fig. 4 also shows distributions of the size of tumor regions, the reconstruction measures $\int_X c_1$, and the fractions ρ_w/κ_w , $vol(TC)/\rho_w$, and $vol(ED)/\kappa_w$, respectively, for each survival class from 161 patients (with survival data). Qualitative and quantitative results for the entire cohort are given in the supplementary.

2) Discussion: Our scheme runs on all scans without any failures and was able to reconstruct well-localized TILs. The model-predicted grown tumor (from estimated growth and infiltration rate) originating from those initial sites, is in good agreement with the observed TC data, as can be seen from Fig. 4. This is quantitatively supported by a small (mean) data-misfit 19% error for μ in Tab. I, with small variations of $\pm 5\%$ across the entire data set. The measure $\int_{ED} c_1$ from Fig. 4 furthermore indicates fair extent of reconstruction of peritumoral edema, while the amount of spurious (false) tumor growth is small with a mean of around 0.1 for $\int_{BWT} c_1$ across all cases.

Regarding the reconstructed TILs, one question that arises is how to evaluate their correctness. Of course, we don't have a ground truth. In [10] we performed tests with synthetic ground truth. Here the unknown biophysical modeling errors and the unknown healthy anatomy of the subjects prevent us from conducting such analysis for the BraTS'18 set. So as a way to estimate the validity of our estimates we compare the reconstructed $\mathbf{p}$ with the center of mass computed using a connected-component analysis of the segmented tumor (see SM). We also check whether our solutions are sparse and localized. Indeed, we observe (i) our estimated $\alpha(0)$ are sparse, and localized, and that (ii) $\mathbf{p}$ is spread out and often quite different from the center of mass of the segmented tumor. This can be seen from the visualization of projected p_i activation (blue dots) in axial, coronal and sagittal views of Fig. 4. On average the center of mass of tumor initiations (per component) deviates from the center of the TC input data by only around 5 voxels. The distance between the TIL activation sites, however, averages at 17 voxels with large outliers (cf. Tab. I and Fig. 4). The former attests plausibility of our solution, while the latter clearly indicates a surplus value of identifying distinct sites of tumor initiation. Lastly, the solution vector $\mathbf{p}$ converges across levels of our coarse-to-fine scheme (cf. projections of red (128^3 and blue (256^3) p_i activation in Fig. 4).

C. Construction of GBM Atlas

We construct a cohort GBM atlas by mapping the reconstructed TILs, $\alpha(0)$, and the predicted grown tumors, $\alpha(1)$, of 161 individuals back into a common atlas space. In a second step, we analyze the distribution of TILs in particular structural and functional regions of the brain. We investigate correlations between probability of tumor occurrence in certain regions with indication towards patient survival. The probabilities are computed per voxel and per region. For a voxel x , we define the probability as $\Sigma_i c_i(x)/\Sigma_{ix} c_i(x)$, where i indicates the case. Similar frequencies are computed for regions by summing all the voxels of a region.

To compute the tumor occurrence frequencies, we register the scans to a healthy template using deformable registration (see §II-B.3). The template image is segmented into structural and functional regions of the brain. We map the TILs, $\alpha(1)$, and segmentation labels to the

atlas space and compute tumor origination frequencies for each region and survival class (short, mid, long).

1) Results: Fig. 1 and Fig. 5 depict the cohort GBM probability atlas. In Fig. 6, we report tumor occurrence probabilities for the most affected brain regions.

2) Discussion: Fig. 1 indicates regions commonly affected by tumors, which lead to overall survival of less than 10 months (*short* survivors; red), between 10 and 15 months (*mid* survivors; green), or more than 15 months (*long* survivors; blue); (see supplementary materials for detailed results). We observe an indication for correlation of short patient survival and tumor origin close to the ventricles (especially lower left ventricle), while tumors originating in white matter closer to the skull correlate with a longer life expectancy. This is in accordance with typically more severe brain damage caused by the tumor mass-effect for tumors originating close to the ventricles [5], [25], [26] and difficulties with resecting such tumors.

In Fig. 6, we see the tumor occurrence probabilities grouped per survival class. Observe that these probabilities differ between the TIL and OT atlases. The OT regional probabilities are not as discriminative as the TIL and are biased due to large tumors from short-term survivors. These large tumors “touch” several regions. Using TIL frequencies allows for more discriminative regional probabilities. For example, see Fig. 6 where TIL shows high frequency for mid (A,F), long (B,D,E), and short (D)—while OT frequencies are uninformative. Of course, due to the quite limited data in this study the results are not definite and suffer from statistical noise. They do suggest, however, issues with using the OT data and the potential benefits of using localized TIL information.

D. Survival Prediction

We conduct a preliminary study for survival prediction for the BraTS’18 survival data. We extract image and physics-based features from 161 individuals⁶ with known overall survival to build a classifier labeling a new subject as either *short*, *mid*, or *long* survivor. We employ XGboost [34], an advanced gradient boosted random forest algorithm for multiclass classification. We consider an 80/20 train/test split, and train the model using 100 rounds, at a learning rate of $\eta = 0.01$ with a maximal tree depth of 5.⁷ The model is trained using the following sets of features:

- A. Patient information:** (i) patient age (A); (ii) treatment status (gross total resection (GTR)/ subtotal resection (STR)/no treatment information);
- B. Image-based features:** (i) volume of *TC*, *ED*, and *NE* relative to brain; (ii) volume fraction $vol(EN)/vol(TC)$ of enhancing tumor relative to tumor core;

⁶short survivors, 41 mid survivors, and 56 long survivors.

⁷The remaining parameters are set to the default values from XGboost’s Python API.

- C. **Physics-based features (tumor inversion):** (i) ratio $\frac{\rho_w}{\kappa_w}$; (ii) ratios $vol(TC)/\rho_w$, and $vol(ED)/\kappa_w$ (relates size of tumor core to cell proliferation, and size of edema to cell infiltration).

1) Results: A classification report on the test set considering the above sets of features is given in Tab. II. Since treatment status is not available for all cases, we also report classification scores for patients which (reportedly) underwent gross total resection (GTR).⁸ Additionally, we report the average accuracy with respect to a five-fold cross-validation on the training data. The determined feature importance is given for each classifier.

2) Discussion: We observe an improvement of all classification scores when augmenting the set of image-based or patient demographics features with additional biophysical features, acquired from our tumor inversion solver. Overall all classifiers generalize reasonably well as the average accuracy from cross-validation compares to the accuracy on the left-out test set.

In Tab. II we also report the feature importance;⁹ we observe that the biophysical features are quite important. In particular, the ratio of proliferation rate to infiltration rate is comparable to the patient's age importance. Due to the significantly smaller dataset when considering cases with GTR only (57 cases),¹⁰ we reduce the number of features to minimize overfitting. The ratio ρ_w/κ_w appears to be most informative. Classification scores deteriorate if further imaging-based features are considered; we explain this by overfitting and a too small dataset.

We emphasize the preliminary nature of this study. The overall classification scores are fairly low, indicating the difficulty of ML methods to handle the large existing heterogeneity of the patient-specific information available, given the limited data. Similar conclusions have been drawn from the official BraTS survival prediction challenge [11], where the winning teams report great variability in classification accuracy depending on the actual set of evaluation samples.

IV. Conclusion

We presented a methodology for reconstructing tumor growth parameters and sparse tumor initiation locations (TILs) from clinical glioblastoma MR imaging data by means of tumor-growth inversion methods. We applied it on 206 subjects from the BraTS'18 training set. We simultaneously estimate (i) the specific tumor cell proliferation rate, (ii) the tumor cell migration rate, and (iii) the original, sparse tumor initial locations. These quantities, combined with image features, were used in a shallow classifier for survival prediction. Also, the TILs and the observed tumor margins were used to construct two different

⁸Treatment status (patients underwent gross total resection) is available for only 57/161 cases.

⁹For a single decision tree, importance is calculated by the amount that each attribute split point improves the performance measure, weighted by the number of observations the node is responsible for. Feature importances are averaged across all of the the decision trees in the model, see [35, p.367]

¹⁰Considering the treatment status as additional feature when using the full dataset instead results in low classification scores; this can be explained by incomplete information compromising the classifier's generalization.

probabilistic atlases of tumor occurrence. We presented preliminary results, relating patient survival to the frequency of tumor occurrence in a specific functional region of the brain.

Some key findings of our study are the following: • Our method can be reliably applied to clinical data in a fully automatic fashion. • TIL localization results in dramatically different infiltration prediction and tumor growth parameters. • The localized biophysical inversion TIL correlates with an entirely image-based TIL computed with connected-component analysis of the tumor segmentation. However, the image-based are quite often different from the TIL and cannot reliably be used to estimate tumor parameters. • The probabilistic tumor occurrence atlas using the TIL differs from the atlas reconstructed using the observed tumor margins. In particular, the latter seems to be biased by the tumor size of particular subjects. • Preliminary results in classification of survival are encouraging but inconclusive due to the small sample size. • We observed that the area around the lower left ventricle (left temporal lobe) shows a clustering of aggressive tumors, while tumors resulting in longer survival are generally found closer to the skull.

Here are some of the *limitations* of our work. We use a phenomenological, single-species model, that cannot distinguish between different glioma phenotypes observed from medical MRI such as necrotic core, enhancing and non-enhancing tumor, and peritumoral edema. We do not account for mass effect; we have derived more complex models that do so [36], and their integration with the sparse TIL reconstruction is ongoing. Our dataset is quite small and voxelwise/regional information is questionable. We have not included comparisons with postoperative recurrence, which is another way to validate the model. (Such postoperative datasets are not available for the BraTS'18 images.) The software is not yet open source but our plan is to make it available through the <https://www.med.upenn.edu/cbica/captk/CaPTk> interface in the near future.

Our method lays the basis for reliable biophysical model calibration from a single preoperative scan. In particular the location of tumor initiation could be used to differentiate tumor mutations, and could (among other radiographic features and biomarkers) contribute towards important clinical diagnostic and prognostic outcomes such as prediction of patient survival times, and identification of brain regions that are of high risk to develop tumors. However, testing these hypotheses conclusively requires further studies with larger datasets and richer biophysical models. Our next steps include the combination of machine learning methods with more complex multispecies tumor models.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

Any opinions, findings, and conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the views of the AFOSR, DOE, NIH, and NSF. Computing time on the Texas Advanced Computing Centers Stampede system was provided by an allocation from TACC and the NSF.

This work was supported in part by the National Institutes of Health (NIH) under Award 5R01NS042645-11A1, in part by the National Science Foundation (NSF) under Award CCF-1817048 and Award CCF-1725743, in part by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied

Mathematics Program under Award DE-SC0019393, in part by the U.S. Air Force Office of Scientific Research under Award FA9550-17-1-0190, and in part by the Portugal Foundation for Science and Technology (FCT) and the UT Austin-Portugal Program.

References

- [1]. Yankeelov T et al. , “Clinically relevant modeling of tumor growth and treatment response,” *Sci. Transl. Med.*, vol. 5, no. 187, p. 187ps9, 2013.
- [2]. Rockne RC et al. , “The 2019 mathematical oncology roadmap,” *Phys. Biol.*, vol. 16, no. 4, p. 041005, 2019. [PubMed: 30991381]
- [3]. Lipkova J et al. , “Personalized radiotherapy design for glioblastoma: Integrating mathematical tumor models, multimodal scans, and Bayesian inference,” *IEEE Trans. Med. Imag.*, vol. 38, no. 8, pp. 1875–1884, 8. 2019.
- [4]. Rathore S et al. , “Radiomic signature of infiltration in peritumoral edema predicts subsequent recurrence in glioblastoma: Implications for personalized radiotherapy planning,” *J. Med. Imag.*, vol. 5, no. 2, 4. 2018, Art. no. 021219.
- [5]. Lee JH and Lee JH, “The origin-of-cell harboring cancer-driving mutations in human glioblastoma,” *BMB Rep.*, vol. 51, no. 10, pp. 481–483, 10. 2018. [PubMed: 30269745]
- [6]. Roux A et al. , “MRI atlas of IDH wild-type supratentorial glioblastoma: Probabilistic maps of phenotype, management, and outcomes,” *Radiology*, vol. 293, 12. 2019, Art. no. 190491.
- [7]. Ren X et al. , “Co-deletion of chromosome 1p/19q and IDH1/2 mutation in glioma subsets of brain tumors in chinese patients,” *PLoS ONE*, vol. 7, no. 3, 3. 2012, Art. no. e32764. [PubMed: 22427879]
- [8]. Bilello M et al. , “Population-based MRI atlases of spatial distribution are specific to patient and tumor characteristics in glioblastoma,” *NeuroImage, Clin.*, vol. 12, pp. 34–40, 2. 2016. [PubMed: 27358767]
- [9]. Li H-Y, Sun C-R, He M, Yin L-C, Du H-G, and Zhang J-M, “Correlation between tumor location and clinical properties of glioblastomas in frontal and temporal lobes,” *World Neurosurgery*, vol. 112, pp. e407–e414, 4. 2018. [PubMed: 29355809]
- [10]. Subramanian S, Scheufele K, Mehl M, and Biros G, “Where did the tumor start? An inverse solver with sparse localization for tumor growth models,” 2019, arXiv:1907.06564. [Online]. Available: <http://arxiv.org/abs/1907.06564>
- [11]. Bakas S et al. , “Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge,” 2018, arXiv:1811.02629. [Online]. Available: <http://arxiv.org/abs/1811.02629>
- [12]. Swanson KR, Rostomily RC, and Alvord EC, “A mathematical modelling tool for predicting survival of individual patients following resection of glioblastoma: A proof of principle,” *Brit. J. Cancer*, vol. 98, no. 1, pp. 113–119, 1. 2008. [PubMed: 18059395]
- [13]. Hawkins-Daarud A, Rockne RC, Anderson ARA, and Swanson KR, “Modeling tumor-associated edema in gliomas during anti-angiogenic therapy and its impact on imageable tumor,” *Frontiers Oncol.*, vol. 3, pp. 1–12, 4. 2013.
- [14]. Clatz O et al. , “Realistic simulation of the 3-D growth of brain tumors in MR images coupling diffusion with biomechanical deformation,” *IEEE Trans. Med. Imag.*, vol. 24, no. 10, pp. 1334–1346, 10. 2005.
- [15]. Hogue C, Davatzikos C, and Biros G, “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, 6. 2008. [PubMed: 18026731]
- [16]. Konukoglu E et al. , “Image guided personalization of reaction-diffusion type tumor growth models using modified anisotropic eikonal equations,” *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 77–95, 1. 2010.
- [17]. Rekik I et al. , “Tumor growth parameters estimation and source localization from a unique time point: Application to low-grade gliomas,” *Comput. Vis. Image Understand.*, vol. 117, no. 3, pp. 238–249, 3. 2013.

- [18]. Jackson PR, Juliano J, Hawkins-Daarud A, Rockne RC, and Swanson KR, "Patient-specific mathematical neuro-oncology: Using a simple proliferation and invasion tumor model to inform clinical practice," *Bull. Math. Biol.*, vol. 77, no. 5, pp. 846–856, 5 2015. [PubMed: 25795318]
- [19]. Gooya A et al. , "GLISTR: Glioma image segmentation and registration," *IEEE Trans. Med. Imag.*, vol. 31, no. 10, pp. 1941–1954, 10. 2012.
- [20]. Macyszyn L et al. , "Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques," *Neuro-Oncol.*, vol. 18, no. 3, pp. 417–425, 3. 2016. [PubMed: 26188015]
- [21]. Scheufele K, Mang A, Gholami A, Davatzikos C, Biros G, and Mehl M, "Coupling brain-tumor biophysical models and diffeomorphic image registration," *Comput. Methods Appl. Mech. Eng.*, vol. 347, pp. 533–567, 4. 2019. [PubMed: 31857736]
- [22]. Menze BH et al. , "A generative approach for image-based modeling of tumor growth," in *Proc Inf. Process. Med. Imag.*, vol. 22, 2011, pp. 735–747.
- [23]. Parisot S et al. , "A probabilistic atlas of diffuse WHO grade II glioma locations in the brain," *PLoS ONE*, vol. 11, no. 1, 1. 2016, Art. no. e0144200. [PubMed: 26751577]
- [24]. Binder ZA et al. , "Epidermal growth factor receptor extracellular domain mutations in glioblastoma present opportunities for clinical imaging and therapeutic development," *Cancer cell*, vol. 34, no. 1, pp. 163–177, 2018. [PubMed: 29990498]
- [25]. Jungk C et al. , "Location-dependent patient outcome and recurrence patterns in IDH1-wildtype glioblastoma," *Cancers*, vol. 11, no. 1, p. 122, 1. 2019.
- [26]. Lee JH et al. , "Human glioblastoma arises from the distant subventricular zone normal appearing but harboring tumor-initiating mutations," *Cancer Res*, vol. 77, no. 13, p. 2455, 2017.
- [27]. Adeberg S, Bostel T, König L, Welzel T, Debus J, and Combs SE, "A comparison of long-term survivors and short-term survivors with glioblastoma, subventricular zone involvement: A predictive factor for survival?" *Radiat. Oncol.*, vol. 9, no. 1, p. 95, 12. 2014. [PubMed: 24758192]
- [28]. Chaichana KL, McGirt MJ, Frazier J, Attenello F, Guerrero-Cazares H, and Quinones-Hinojosa A, "Relationship of glioblastoma multiforme to the lateral ventricles predicts survival following tumor resection," *J. Neuro-Oncol.*, vol. 89, no. 2, pp. 219–224, 9. 2008.
- [29]. Murray JD, *Mathematical Biology*. New York, NY, USA: Springer, 1989.
- [30]. Gholami A et al., "A novel domain adaptation framework for medical image segmentation," in *Proc. Int. MICCAI Brainlesion Workshop*. Cham, Switzerland: Springer, 9. 2018, pp. 289–298.
- [31]. Gholami A, Mang A, and Biros G, "An inverse problem formulation for parameter estimation of a reaction–diffusion model of low grade gliomas," *J. Math. Biol.*, vol. 72, nos. 1–2, pp. 409–433, 1. 2016. [PubMed: 25963601]
- [32]. Needell D and Tropp JA, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 5 2009.
- [33]. Mang A, Gholami A, Davatzikos C, and Biros G, "CLAIRE: A distributed-memory solver for constrained large deformation diffeomorphic image registration," *SIAM J. Sci. Comput.*, vol. 41, no. 5, pp. C548–C584, 1. 2019. [PubMed: 34650324]
- [34]. Chen T and Guestrin C, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*. New York, NY, USA: ACM, 8. 2016, pp. 785–794.
- [35]. Hastie T, Tibshirani R, and Friedman J, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2009.
- [36]. Subramanian S, Gholami A, and Biros G, "Simulation of glioblastoma growth using a 3D multispecies tumor model with mass effect," *J. Math. Biol.*, vol. 79, no. 3, pp. 941–967, 5 2019. [PubMed: 31127329]

Fig. 1.

Cohort GBM statistical atlas for BratTS'18 data. Top figures: Atlas showing relative frequency of tumor initiation location (TIL) (left, subfigures a–c), and observed tumor (OT) (right, d–f) for the Brats'18 survival data (161 cases; color coded by survival class: short survivors red; mid survivors green; long survivors blue). Bottom: Probabilities for TIL (g–i) and OT (j–l) across entire cohort of 206 cases (TIL max probability is 0.007; OT max probability is 0.1). Applying our algorithm to larger datasets and more accurate models could help test a recent hypothesis that certain types of GBMs originate in the subventricular zone [5].

Fig. 2.

Comparison of localized (sparse) and non-localized TIL reconstruction and tumor prediction. Each row corresponds to a different BraTS subject. The first three columns are localized TIL; the last three are non-localized TIL. The observed tumor (OT) segmentation that drives the inversion is the white (dashed) contour line. Subfigures (a) and (d) show the observed tumor (OT) reconstruction at the scan time ($t=1$) as the solution of a tumor growth problem colored from red (100% tumor saturation) to violet (5% saturation). The TIL is in gray and contained in the support of OT. We see that the predictions of the observed tumor concentration at two future times ($t=1.2$ and $t=1.5$) differ a lot when comparing the localized (c–d) and non-localized (e–f) TIL reconstructions. The non-localized TIL underpredicts infiltration.

Fig. 3.

Sensitivity to observation operator. *Qualitative results:* For $\lambda = 1$ (panels a); d)), we assume no observation in edema (i.e., this region is masked from the data misfit term); for $\lambda = 0.1$ (panels b); e)), we additionally discount data mismatch outside the whole tumor by a factor of 0.1; for “observation everywhere” (panels c); f)), no observation operator is considered. The model prediction $c(1)$ is given in contours from red to violet; the reconstructed TIL, $c(0)$, is shown in dark gray; the TC input data is indicated as a dashed white line. Additionally, a one-dimensional representation of reconstructed tumor cell densities across the axial-x (for TCIA02_296) and axial-y (for CBICA_AXO) line is overlaid with the DNN segmentation of TC and ED. Tumor cell density reconstructions are given in red for $\lambda = 1$, in green for $\lambda = 0.1$, and in blue for observation everywhere. *Quantitative results:* Calibrated parameters ρ_w and κ_w , and data misfit μ for different the same observation operators (“-” indicates observation everywhere). Additionally, we report the misfit μ_B in the entire brain, and the integral measures $\int_{ED} c_1 = \int_{ED} c(1) dx / \int_{\Omega} c(1) dx$ (approximation of edema), and $\int_{B \setminus WT} c_1 = \int_{B \setminus WT} c(1) dx / \int_{\Omega} c(1) dx$ (tumor prediction in healthy tissue). More results can be found in the supplementary material.

Fig. 4.

BraTS'18 Results. ■ Left: Shown are axial, coronal, and sagittal 3D views of the predicted tumor (blue), its estimated origin (red), and CSF with ventricles (green). Right: Projection of estimated TIL p onto axial, coronal, and sagittal plane (red corresponds to a 1283 image resolution; blue corresponds to a 2563 resolution). The size of the markers indicates the magnitude of p . The center of mass of tumor core input data components are indicated with a black cross. ■ **Multifocal Cases.** The slices are selected at coordinates of the center of mass of the largest (left) and second largest (middle) TC component. The biophysical model-predicted tumor $\alpha(1)$ is given in contours from red to violet; the reconstructed TIL, $\alpha(0)$, is shown in dark gray and is always in the support of $\alpha(1)$; the TC input data is indicated as a dashed white line. The red $\alpha(1)$ contour line corresponds to $0.9 \cdot \alpha(1)$, and matches well with the input TC data. The right-hand side shows TIL projections as explained above. We refer to multifocal cases as tumors featuring at least two disconnected components of the TC ; such TC components may either be connected by FLAIR abnormality, or represent multicentric tumors. ■ **Single-focal Cases.** ■ **Cohort Statistics.** Depicted are median and variance of relative size, and reconstruction measures $\int_X c_1 = \int_X c(1) dx / \int_{\Omega} c(1) dx$, with $X \in \{TC, ED, B \setminus WT\}$ of tumor subregions, subdivided into three survival classes (long, mid, short). We also report the distribution of ρ_w / κ_w , $vol(TC) / \rho_w$, and $vol(ED) / \kappa_w$ for the three classes. Boxes range from lower to upper quartile.

Fig. 5.

Cohort GBM atlas (GBM-A) for BratTS'18 survival data. *Top*: Probability distributions of observed tumor (OT) and tumor Initiation Locations (TILs) of 161 patients superimposed and displayed in the atlas space (given for the entire cohort, short, mid, and long survivors). *Middle*: TILs of entire cohort displayed in the atlas space with short survivors in red, mid survivors in green, and long survivors in blue. *Bottom*: Frequently affected functional regions.

Fig. 6.

Correlation of patient survival with TIL in certain functional regions based on the probabilistic GBM atlas a (cf. Fig. 1), we compute TIL and OT frequencies $\int_{\mathcal{L}} a_X(x) dx / \int_{\Omega} a(x) dx$ per functional region $\mathcal{L}$; a_X for $X \in \{S, M, L\}$ denotes the probabilistic atlas of short, mid, and long survivors. We report the ratio of (TIL/OT) frequencies for short (blue), mid (green), and long (cyan) survivors per label (imbalances indicate higher likelihood for survival class in this region), along with the rel. size and overall percentage (summed across short, mid, long) of TIL ($f(c_0)$) and OT ($f(c_1)$). We display the top 16 most frequently affected regions of 145 total (see supplementary materials for detailed results).

TABLE I:

Top: Quantitative Results. We Report the Estimated Tumor Proliferation and Infiltration Rates ρ_w , and κ_w , the Relative ℓ_2 -Data Misfit at Observation Points μ , the Distance D_{cm} (in Voxels) Between the Center of Mass (CM) of Tumor Initiation Sites $\alpha(0)$ and the CM of TC (of the Largest Component), and the Distance D_p (in Voxels) Between the Two Most Important Tumor Initiation Sites. Mean and Standard Deviation Are Given for the Entire BraTS'18 cohort of 206 Cases. **Bottom:** Median and Variance of the Misfit μ , and Voxel Distances D_{cm} and D_p . Boxes Range From Lower to Upper Quartile

BraTS ID	ρ_w	κ_w	μ	D_{cm}	D_p
CBICA_ASQ (2)	11.1	1.53E-2	1.73E-1	2.8	19.6
CBICA_AOZ (3)	9.49	9.96E-3	1.92E-1	3.9	19.6
CBICA_AQU (2)	9.66	1.14E-2	1.88E-1	1.3	25.3
CBICA_AME (3)	10.1	1.04E-2	1.93E-1	3.6	19.6
TCIA02_314 (1)	9.7	1.07E-2	2.62E-1	4.1	49.3
TCIA03_257 (2)	9.2	1.11E-2	2.19E-1	4.2	17.8
TCIA05_277 (1)	10.9	2.49E-2	1.64E-1	8.1	51.2
TCIA03_296 (1)	11.6	2.76E-2	1.80E-1	7.1	19.6
TCIA02_606 (1)	11.6	2.68E-2	1.74E-1	5.4	40.0
TCIA01_180 (1)	12.7	1.39E-2	1.37E-1	6.7	17.8
CBICA_BHM (1)	9.17	9.31E-3	1.67E-1	5.5	16.0
<i>Mean</i>	1.03E1	1.37E-2	1.90E-1	3.99E0	1.73E1
<i>Std. Deviation</i>	2.65E0	9.75E-3	5.52E-2	2.98E0	8.72E0

TABLE II:

Classification of Patient Survival for Image-Based, and Biophysically Augmented Features. The Total Number of Samples is 161 (64 Short, 41 Mid, and 56 Long Survivors); 57/161 Cases (20 Short, 14 Mid, 23 Long) Underwent Gross Total Resection (GTR). Classification Results Are Given for the Entire Dataset, and for Cases With GTR Only; Scores Are Reported as Weighted Mean. We Use an 80/20 Train/Test Split, and Also Report the Average Accuracy for a 5-Fold CV. Feature Importance (and Use) is Given for Each Classifier. $\frac{vol(TC)}{\rho}$ Has Been Removed From “A+Img.+Phys.” Due To Colinearity and High VIF (See Supplementary Material)

Features	Training Data				Test Data				Feature Importance			
	CV avg.-accuracy	accuracy	precision	recall	β -score	$\frac{\rho W}{KW}$	Age	$\frac{vol(TC)}{\rho W}$	$\frac{vol(TC)}{KW}$	$\frac{vol(ED)}{KW}$	$\frac{vol(EN)}{vol(TC)}$	$\frac{vol(ED)}{vol(TC)}$
<i>Full-Dataset</i>	Age (A)	0.43±0.11	0.33	0.33	0.34	0.33	1.00	–	–	–	–	–
	A+Img.	0.46±0.18	0.39	0.37	0.39	0.37	0.28	–	0.28	–	0.21	0.20
	A+Phys.	0.46±0.15	0.46	0.45	0.46	0.45	0.27	0.23	–	0.21	–	–
	A+Img.+Phys.	0.52±0.12	0.51	0.50	0.51	0.49	0.20	–	0.15	0.16	0.14	0.14
<i>GTR Only</i>	Age (A)	0.42±0.24	0.40	0.38	0.40	0.37	1.00	–	–	–	–	–
	A+Img.	0.37±0.09	0.33	0.26	0.33	0.29	0.28	–	0.28	–	0.21	0.20
	A+Phys.	0.56 ±0.11	0.43	0.40	0.43	0.40	0.42	–	–	–	–	–